

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 14:47:14 ; Search time 153 Seconds
(without alignments)
2203.281 Million cell updates/sec

Title: US-10-781-294-24

Perfect score: 5472
Sequence: 1 MLRTAGDGLCRSLTYLEEL.....MTHSLAALRVKPYLDIGC 1035

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1373511 seqs, 325702437 residues

Total number of hits satisfying chosen parameters: 1373511

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppa/PCT05_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubppa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------------------------|
| 1 | 5472 | 100.0 | 1035 | 10 | US-09-965-621-24 Sequence 24, Appl |
| 2 | 5472 | 100.0 | 1035 | 15 | US-10-407-866-24 Sequence 24, Appl |
| 3 | 5472 | 100.0 | 1035 | 16 | US-10-781-294-24 Sequence 24, Appl |
| 4 | 5433 | 99.3 | 1061 | 14 | US-10-124-498-18 Sequence 18, Appl |
| 5 | 5433 | 99.3 | 1061 | 14 | US-10-066-521-18 Sequence 18, Appl |
| 6 | 5259.5 | 96.1 | 1027 | 15 | US-10-407-866-68 Sequence 68, Appl |
| 7 | 5077.5 | 92.8 | 1004 | 15 | US-10-108-260A-3161 Sequence 3161, Ap |
| 8 | 3364 | 61.5 | 635 | 15 | US-10-407-866-90 Sequence 90, Appl |
| 9 | 3163 | 57.8 | 603 | 15 | US-09-965-621-64 Sequence 64, Appl |
| 10 | 3163 | 57.8 | 603 | 15 | US-10-407-866-64 Sequence 64, Appl |
| 11 | 3163 | 57.8 | 603 | 16 | US-10-781-294-64 Sequence 22, Appl |
| 12 | 3005 | 54.9 | 582 | 15 | US-10-162-335-22 Sequence 22, Appl |
| 13 | 2953.5 | 54.0 | 565 | 16 | US-10-467-397-9 Sequence 9, Appl |

| | | | | | |
|----|--------|------|------|----|-------------------------------------|
| 14 | 2694.5 | 49.2 | 521 | 15 | US-10-162-335-24 Sequence 24, Appl |
| 15 | 2396 | 43.8 | 1034 | 13 | US-10-127-516-5 Sequence 5, Appl |
| 16 | 2396 | 43.8 | 1034 | 13 | US-10-027-629-5 Sequence 5, Appl |
| 17 | 2396 | 43.8 | 1034 | 14 | US-10-028-374-18 Sequence 18, Appl |
| 18 | 2396 | 43.8 | 1034 | 14 | US-10-132-967-5 Sequence 5, Appl |
| 19 | 2396 | 43.8 | 1034 | 14 | US-10-183-770-18 Sequence 18, Appl |
| 20 | 2396 | 43.8 | 1034 | 15 | US-10-264-958B-22 Sequence 22, Appl |
| 21 | 2334 | 42.7 | 449 | 14 | US-10-028-374-2 Sequence 2, Appl |
| 22 | 2334 | 42.7 | 449 | 14 | US-10-183-770-2 Sequence 2, Appl |
| 23 | 1749 | 32.0 | 674 | 15 | US-10-407-866-89 Sequence 89, Appl |
| 24 | 1731 | 31.6 | 321 | 15 | US-10-407-866-116 Sequence 116, App |
| 25 | 1664 | 30.2 | 344 | 9 | US-09-799-983-2 Sequence 2, Appl |
| 26 | 1653.5 | 30.2 | 719 | 10 | US-09-965-621-59 Sequence 59, Appl |
| 27 | 1653.5 | 30.2 | 719 | 15 | US-10-407-866-59 Sequence 59, Appl |
| 28 | 1653.5 | 30.2 | 719 | 16 | US-10-781-294-59 Sequence 59, Appl |
| 29 | 1631.5 | 29.8 | 612 | 15 | US-10-407-866-88 Sequence 88, Appl |
| 30 | 1598 | 29.2 | 924 | 15 | US-10-357-820-52 Sequence 52, Appl |
| 31 | 1457.5 | 26.6 | 994 | 10 | US-09-965-621-16 Sequence 16, Appl |
| 32 | 1457.5 | 26.6 | 994 | 14 | US-10-124-498-24 Sequence 24, Appl |
| 33 | 1457.5 | 26.6 | 994 | 14 | US-10-066-521-24 Sequence 24, Appl |
| 34 | 1457.5 | 26.6 | 994 | 15 | US-10-407-866-16 Sequence 16, Appl |
| 35 | 1457.5 | 26.6 | 994 | 16 | US-10-781-294-16 Sequence 16, Appl |
| 36 | 1456.5 | 26.6 | 920 | 16 | US-10-250-615-13 Sequence 13, Appl |
| 37 | 1432.5 | 26.2 | 896 | 14 | US-10-124-498-22 Sequence 22, Appl |
| 38 | 1432.5 | 26.2 | 896 | 14 | US-10-066-521-22 Sequence 22, Appl |
| 39 | 1385 | 25.3 | 1473 | 15 | US-10-182-822A-8 Sequence 8, Appl |
| 40 | 1381 | 25.2 | 1429 | 14 | US-10-029-347-3 Sequence 3, Appl |
| 41 | 1381 | 25.2 | 1429 | 14 | US-10-028-374-3 Sequence 3, Appl |
| 42 | 1381 | 25.2 | 1429 | 14 | US-10-183-770-3 Sequence 3, Appl |
| 43 | 1375 | 25.1 | 1429 | 9 | US-09-996-617-2 Sequence 2, Appl |
| 44 | 1375 | 25.1 | 1429 | 9 | US-09-931-071-2 Sequence 2, Appl |
| 45 | 1375 | 25.1 | 1429 | 14 | US-10-029-347-26 Sequence 26, Appl |

ALIGNMENTS

RESULT 1
US-09-965-621-24
; Sequence 24, Application US/09965621
; Publication No. US20030077699A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/09/965,621
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1035
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-965-621-24

| | | | | |
|-----------------------|-----------------|---|-----------|--------------|
| Query Match | 100.0%; | Score 5472; | DB 10; | Length 1035; |
| Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| Matches 1035; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 1 | MLRTAGDGLCRSLTYLEELAAVELKKFKLYLGTATELGEGKIPWGSMEKAGPLEMAQLL | 60 | |
| Db | 1 | MLRTAGDGLCRSLTYLEELAAVELKKFKLYLGTATELGEGKIPWGSMEKAGPLEMAQLL | 60 | |
| Qy | 61 | ITHFGPEAWRLALSTFERINRKDLWREGQDLVRDPOETRYDYVRRKFLMEDRNARL | 120 | |

| | | | | | |
|----|-----|---|-----|-----|--|
| | | ; APPLICANT: Reed, John C. | | | |
| | | ; APPLICANT: Godzik, Adam | | | |
| | | ; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides, | | | |
| | | ; FILE REFERENCE: 66654-10 (LJ 5755) | | | |
| | | ; CURRENT APPLICATION NUMBER: US/10/407,866 | | | |
| | | ; CURRENT FILING DATE: 2003-04-04 | | | |
| | | ; PRIOR APPLICATION NUMBER: US 60/370,538 | | | |
| | | ; PRIOR FILING DATE: 2002-04-04 | | | |
| | | ; NUMBER OF SEQ ID NOS: 129 | | | |
| | | ; SOFTWARE: FastSeq for Windows Version 4.0 | | | |
| | | ; SEQ ID NO 24 | | | |
| | | ; LENGTH: 1035 | | | |
| | | ; TYPE: PRT | | | |
| | | ; ORGANISM: Homo sapien | | | |
| | | ; US-10-407-866-24 | | | |
| | | Query Match 100.0%; Score 5472; DB 15; Length 1035; | | | |
| | | Best Local Similarity 100.0%; Pred. No. 0; | | | |
| | | Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | |
| Db | 61 | ITHFGPEEAWRLALSTFERINRKDLWERGQREDLVDPQETVRDYVRRKFRMLMEDRNARL | 120 | 1 | MLRTAGRDGLCRSLSTYLEEAEVLEKFKLYLTATLGEKGIPWGSMEKAGPLEMAQLL |
| Qy | 121 | GEVCNLSHRYTRLLLVKHSNPMQVQQLDTRGHARTVGHQASPIKTIETLFEDEBRP | 180 | 1 | MLRTAGRDGLCRSLSTYLEEAEVLEKFKLYLTATLGEKGIPWGSMEKAGPLEMAQLL |
| Db | 121 | GEVCNLSHRYTRLLLVKHSNPMQVQQLDTRGHARTVGHQASPIKTIETLFEDEBRP | 180 | 61 | ITHFGPEEAWRLALSTFERINRKDLWERGQREDLVDPQETVRDYVRRKFRMLMEDRNARL |
| Qy | 181 | EPRTVMQGAAGIGKSMIAHKVMDWADGKLPQGRFDYLFYINCREMNSQATECSMDL | 240 | 61 | ITHFGPEEAWRLALSTFERINRKDLWERGQREDLVDPQETVRDYVRRKFRMLMEDRNARL |
| Db | 181 | EPRTVMQGAAGIGKSMIAHKVMDWADGKLPQGRFDYLFYINCREMNSQATECSMDL | 240 | 121 | GEVCNLSHRYTRLLLVKHSNPMQVQQLDTRGHARTVGHQASPIKTIETLFEDEBRP |
| Qy | 241 | IFSCWPEPSAPLOELIRVPERLLFTIDGDELKPSFHDPOGFWCLCWEKRPTELLNSL | 300 | 121 | GEVCNLSHRYTRLLLVKHSNPMQVQQLDTRGHARTVGHQASPIKTIETLFEDEBRP |
| Db | 241 | IFSCWPEPSAPLOELIRVPERLLFTIDGDELKPSFHDPOGFWCLCWEKRPTELLNSL | 300 | 181 | EPRTVMQGAAGIGKSMIAHKVMDWADGKLPQGRFDYLFYINCREMNSQATECSMDL |
| Qy | 301 | IRKLLPELSLITTRPTALEKHLRLEHPRHVEILGFSEAEKKEYFYKFNABQAGV | 360 | 181 | EPRTVMQGAAGIGKSMIAHKVMDWADGKLPQGRFDYLFYINCREMNSQATECSMDL |
| Db | 301 | IRKLLPELSLITTRPTALEKHLRLEHPRHVEILGFSEAEKKEYFYKFNABQAGV | 360 | 241 | IFSCWPEPSAPLOELIRVPERLLFTIDGDELKPSFHDPOGFWCLCWEKRPTELLNSL |
| Qy | 361 | FNVRDNEPLFTMCFVPLVCVVCTCLOQLEGGLLROTSTRTTAVYMLYLLSLMQPK | 420 | 241 | IFSCWPEPSAPLOELIRVPERLLFTIDGDELKPSFHDPOGFWCLCWEKRPTELLNSL |
| Db | 361 | FNVRDNEPLFTMCFVPLVCVVCTCLOQLEGGLLROTSTRTTAVYMLYLLSLMQPK | 420 | 301 | IRKLLPELSLITTRPTALEKHLRLEHPRHVEILGFSEAEKKEYFYKFNABQAGV |
| Qy | 421 | GAPRLQPPNQRGLCSLAADGLWNQKILFEODLRKHGLDGEDVSAFLNWNIFQKDINCE | 480 | 301 | IRKLLPELSLITTRPTALEKHLRLEHPRHVEILGFSEAEKKEYFYKFNABQAGV |
| Db | 421 | GAPRLQPPNQRGLCSLAADGLWNQKILFEODLRKHGLDGEDVSAFLNWNIFQKDINCE | 480 | 361 | FNVRDNEPLFTMCFVPLVCVVCTCLOQLEGGLLROTSTRTTAVYMLYLLSLMQPK |
| Qy | 481 | RYYSFIHLSFOEFPAAMYIILDEGEGAGPDQDVTRLLTEYAFSERSFLATSRFLGGL | 540 | 361 | FNVRDNEPLFTMCFVPLVCVVCTCLOQLEGGLLROTSTRTTAVYMLYLLSLMQPK |
| Db | 481 | RYYSFIHLSFOEFPAAMYIILDEGEGAGPDQDVTRLLTEYAFSERSFLATSRFLGGL | 540 | 421 | GAPRLQPPNQRGLCSLAADGLWNQKILFEODLRKHGLDGEDVSAFLNWNIFQKDINCE |
| Qy | 541 | NEETRSHLEKSLCWKVSPIKMDLQWIOSKASDGSSTLQOQSLFSPCLYEIOEEFIQ | 600 | 421 | GAPRLQPPNQRGLCSLAADGLWNQKILFEODLRKHGLDGEDVSAFLNWNIFQKDINCE |
| Db | 541 | NEETRSHLEKSLCWKVSPIKMDLQWIOSKASDGSSTLQOQSLFSPCLYEIOEEFIQ | 600 | 481 | RYYSFIHLSFOEFPAAMYIILDEGEGAGPDQDVTRLLTEYAFSERSFLATSRFLGGL |
| Qy | 601 | QALSHFOVIVVSNIAKMEHWSSFCILKRCRQAQVLLHYGATYSADGEDRARCAGAHTL | 660 | 481 | RYYSFIHLSFOEFPAAMYIILDEGEGAGPDQDVTRLLTEYAFSERSFLATSRFLGGL |
| Db | 601 | QALSHFOVIVVSNIAKMEHWSSFCILKRCRQAQVLLHYGATYSADGEDRARCAGAHTL | 660 | 541 | NEETRSHLEKSLCWKVSPIKMDLQWIOSKASDGSSTLQOQSLFSPCLYEIOEEFIQ |
| Qy | 661 | LVQLRPRTVLLDAYSEHLAAALCTPNLIELSLYRNALSGRGVYLLCOGLRHPNCKLQ | 720 | 541 | NEETRSHLEKSLCWKVSPIKMDLQWIOSKASDGSSTLQOQSLFSPCLYEIOEEFIQ |
| Db | 661 | LVQLRPRTVLLDAYSEHLAAALCTPNLIELSLYRNALSGRGVYLLCOGLRHPNCKLQ | 720 | 601 | QALSHFOVIVVSNIAKMEHWSSFCILKRCRQAQVLLHYGATYSADGEDRARCAGAHTL |
| Qy | 721 | LRLKRCRISSACEDLSAALIANKNLTRMDLSNGVGFPGCMMLCEGLRHPQCRLOMIQL | 780 | 601 | QALSHFOVIVVSNIAKMEHWSSFCILKRCRQAQVLLHYGATYSADGEDRARCAGAHTL |
| Db | 721 | LRLKRCRISSACEDLSAALIANKNLTRMDLSNGVGFPGCMMLCEGLRHPQCRLOMIQL | 780 | 661 | LVQLRPRTVLLDAYSEHLAAALCTPNLIELSLYRNALSGRGVYLLCOGLRHPNCKLQ |
| | | RESULT 2 | | | |
| | | US-10-407-866-24 | | | |
| | | ; Sequence 24, Application US/10407866 | | | |
| | | ; Publication No. US20040002593A1 | | | |
| | | ; GENERAL INFORMATION: | | | |

Qy 781 RKQLESGACQEMASVLTGTPHVLVELDITGNALBDLGLRLLCQGLRHPVCRRLTLWLKIC 840
Db 781 RKQLESGACQEMASVLTGTPHVLVELDITGNALBDLGLRLLCQGLRHPVCRRLTLWLKIC 840
Qy 841 RLTAACDELASTLSVNSQSLRELDLSNELDGLGVLLICEGLRHPTCKLQTLRLGICRLG 900
Db 841 RLTAACDELASTLSVNSQSLRELDLSNELDGLGVLLICEGLRHPTCKLQTLRLGICRLG 900
Qy 901 SAACEGLSVVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLDSGLTAKA 960
Db 901 SAACEGLSVVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLDSGLTAKA 960
Qy 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLLCRLSHPGCKRLVLMFGMDLNKMTSHR 1020
Db 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLLCRLSHPGCKRLVLMFGMDLNKMTSHR 1020
Qy 1021 LAALRVTKPYLDIGC 1035
Db 1021 LAALRVTKPYLDIGC 1035

RESULT 3
US-10-781-294-24
; Sequence 24, Application US/10781294
; Publication No. US20040142374A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/10781,294
; PRIOR FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1035
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-781-294-24

Query Match 100.0%; Score 5472; DB 16; Length 1035;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLRTAGRDGLCRSLTYLEELAEVLEKPKLYLTATLGECKIPWGSMEKAGPLEMAQLL 60
Db 1 MLRTAGRDGLCRSLTYLEELAEVLEKPKLYLTATLGECKIPWGSMEKAGPLEMAQLL 60
Qy 61 ITHFGPEAWFLALSTFERINRKNLWERGQEDLVDPQETRYDYVRKFLMEDNRAL 120
Db 61 ITHFGPEAWFLALSTFERINRKNLWERGQEDLVDPQETRYDYVRKFLMEDNRAL 120
Qy 121 GECVNLSHRYTRLLLVKEHSPMVQOQLDITGRGHARTVGHQASPIKTIETLPFPDEERP 180
Db 121 GECVNLSHRYTRLLLVKEHSPMVQOQLDITGRGHARTVGHQASPIKTIETLPFPDEERP 180
Qy 181 EPPRTVMQGAAGIGKSLAHKVMLDWADGKLFQGRFDYLYINCREMNQSAATECSMDL 240
Db 181 EPPRTVMQGAAGIGKSLAHKVMLDWADGKLFQGRFDYLYINCREMNQSAATECSMDL 240
Qy 241 IFSCWPEPSAPLOELIRVPERLLFIIDGFDLKPSPHDQPGWCLCWEEKRPTLELLNSL 300
Db 241 IFSCWPEPSAPLOELIRVPERLLFIIDGFDLKPSPHDQPGWCLCWEEKRPTLELLNSL 300

Qy 301 IRKLLPELSLITRPTALEKHLRLLEHPRHVEILGFSEAEKEYFYKYFHNAEQGV 360
Db 301 IRKLLPELSLITRPTALEKHLRLLEHPRHVEILGFSEAEKEYFYKYFHNAEQGV 360
Qy 361 FNYVRDNEPLFTMCVFPLVGVVCTCLOQLEGGGLLRQTSRTTAVYMLYLLSLMQPKP 420
Db 361 FNYVRDNEPLFTMCVFPLVGVVCTCLOQLEGGGLLRQTSRTTAVYMLYLLSLMQPKP 420
Qy 421 GAPRLQPPPNORGICSLAADGLWNQKILFEEODLRKHGLDGEDVSALNNMIFOKDINCE 480
Db 421 GAPRLQPPPNORGICSLAADGLWNQKILFEEODLRKHGLDGEDVSALNNMIFOKDINCE 480
Qy 481 RYYSFIHLSFQEPFAAMYIILDEGEAGAGPDQDVTLLTEYAFSERSFLALTSRFLGGL 540
Db 481 RYYSFIHLSFQEPFAAMYIILDEGEAGAGPDQDVTLLTEYAFSERSFLALTSRFLGGL 540
Qy 541 NEETRSHLEKSLCWKVSFPHIKMDLLQWIOSKAQSDGSTLOQGSLEFSCLYEIEEFIQ 600
Db 541 NEETRSHLEKSLCWKVSFPHIKMDLLQWIOSKAQSDGSTLOQGSLEFSCLYEIEEFIQ 600
Qy 601 QALSHFOVIVVSNIASKMEHVMVSSFCVKRCRSQVTLHLYGATYSADGEDRARCAGATL 660
Db 601 QALSHFOVIVVSNIASKMEHVMVSSFCVKRCRSQVTLHLYGATYSADGEDRARCAGATL 660
Qy 661 LVQLRPBTVLLDAYSEHLAAALCTNPNLIELSLYRNALSGRGVKKLCCQGLRHPNCKLQN 720
Db 661 LVQLRPBTVLLDAYSEHLAAALCTNPNLIELSLYRNALSGRGVKKLCCQGLRHPNCKLQN 720
Qy 721 LRLKRCRISSACEDLSAALIANKVLRMDLSGNGVFPGMMLLCEGLRHPQCRLOMTQL 780
Db 721 LRLKRCRISSACEDLSAALIANKVLRMDLSGNGVFPGMMLLCEGLRHPQCRLOMTQL 780
Qy 781 RKQLESGACQEMASVLTGTPHVLVELDITGNALBDLGLRLLCQGLRHPVCRRLTLWLKIC 840
Db 781 RKQLESGACQEMASVLTGTPHVLVELDITGNALBDLGLRLLCQGLRHPVCRRLTLWLKIC 840
Qy 841 RLTAACDELASTLSVNSQSLRELDLSNELDGLGVLLICEGLRHPTCKLQTLRLGICRLG 900
Db 841 RLTAACDELASTLSVNSQSLRELDLSNELDGLGVLLICEGLRHPTCKLQTLRLGICRLG 900
Qy 901 SAACEGLSVVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLDSGLTAKA 960
Db 901 SAACEGLSVVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLDSGLTAKA 960
Qy 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLLCRLSHPGCKRLVLMFGMDLNKMTSHR 1020
Db 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLLCRLSHPGCKRLVLMFGMDLNKMTSHR 1020
Qy 1021 LAALRVTKPYLDIGC 1035
Db 1021 LAALRVTKPYLDIGC 1035

RESULT 4
US-10-124-498-18
; Sequence 18, Application US/10124498
; Publication No. US20030017983A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Wang, Weiye
; APPLICANT: Blatcher, Maria
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-367001
; CURRENT APPLICATION NUMBER: US/10/124,498
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 10/066,521
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/318,645
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/265,231
; PRIOR FILING DATE: 2001-01-31

```

; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1061
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-124-498-18

Query Match      99.3%; Score 5433; DB 14; Length 1061;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1034; Conservative 0; Mismatches 0; Indels 28; Gaps 2;

QY 1 MLRTAGRDGICRLSTYLELEAVELEKFKLYLGATATGEGKIPWGSMEKAGLEMAQLL 60
DB 1 MLRTAGRDGICRLSTYLELEAVELEKFKLYLGATATGEGKIPWGSMEKAGLEMAQLL 60
QY 61 ITHFGPEAWRLALSTFERINRKLWERGQREDLVR----- 96
DB 61 ITHFGPEAWRLALSTFERINRKLWERGQREDLVR----- 96
QY 97 ---DPOETRYDYVRKFRMLMEDNARLGEVCNLSHRYTRLLLVKHSNPMQVQQLDGTG 153
DB 121 PRKDPQSTYDYVRKFRMLMEDNARLGEVCNLSHRYTRLLLVKHSNPMQVQQLDGTG 180
QY 154 RGHARTVGHQASPIKIETLPEPDEERPEPPRTVVMQGAAGIGKSMIAHKVMDWDGKLF 213
DB 181 RGHARTVGHQASPIKIETLPEPDEERPEPPRTVVMQGAAGIGKSMIAHKVMDWDGKLF 240
QY 214 QGRFDYLYNCREMNSQATECSMDLI FSCWPEPSAPLOELIRVPERLLFIIDGDELK 273
DB 241 QGRFDYLYNCREMNSQATECSMDLI FSCWPEPSAPLOELIRVPERLLFIIDGDELK 300
QY 274 PSFHDPPQPCWCLCWEERKPTLELLNSLIRKLLPELSLITTRPTALEKHLRLEHPRHV 333
DB 301 PSFHDPPQPCWCLCWEERKPTLELLNSLIRKLLPELSLITTRPTALEKHLRLEHPRHV 360
QY 334 EILGFSAEARKEYFYKYFHNAEQAGQVFNVRDNEPLFTWCFVPLVCWVCTCLOOLEG 393
DB 361 EILGFSAEARKEYFYKYFHNAEQAGQVFNVRDNEPLFTWCFVPLVCWVCTCLOOLEG 420
QY 394 GGLLRQTSRTTAVYMLYLLSLMOPKPGAPLQPPNORGLCSLAADGLWNKILFEEQD 453
DB 421 GGLLRQTSRTTAVYMLYLLSLMOPKPGAPLQPPNORGLCSLAADGLWNKILFEEQD 480
QY 454 LRKHGLDGEDVSFAFLNMIFQKIDNCERYYSFIHLSFOEFAAMYIILDEGEGAGDQD 513
DB 481 LRKHGLDGEDVSFAFLNMIFQKIDNCERYYSFIHLSFOEFAAMYIILDEGEGAGDQD 540
QY 514 VTRLLTEYAFSERSFALTSRFLFGLLNEETRSHLEKSLCWKVSPIHKMDLLOWIQSKAQ 573
DB 541 VTRLLTEYAFSERSFALTSRFLFGLLNEETRSHLEKSLCWKVSPIHKMDLLOWIQSKAQ 600
QY 574 SDGSTLQOQSLEFFSCLYEIOBEETIQALSHFQVIVVSNIASKMEHVSSFCLEKCRSA 633
DB 601 SDGSTLQOQSLEFFSCLYEIOBEETIQALSHFQVIVVSNIASKMEHVSSFCLEKCRSA 660
QY 634 QVHLHYGATYSADGEDRARCAGATLLVQLPERTVLLDAYSEHLAALCTNPNLIELS 693
DB 661 QVHLHYGATYSADGEDRARCAGATLLVQLPERTVLLDAYSEHLAALCTNPNLIELS 719
QY 694 LYRNALGSRGVKLLCCOGLRHPNCKLQNLRLKRCRISACEDLSAALIANKNLITRMDLSG 753
DB 720 LYRNALGSRGVKLLCCOGLRHPNCKLQNLRLKRCRISACEDLSAALIANKNLITRMDLSG 779
QY 754 NGVGFPGMMLLCEGLRHPQRLQMIQIRKCOLESAGCQEWASVLTNPHLVELDITGNAL 813
DB 780 NGVGFPGMMLLCEGLRHPQRLQMIQIRKCOLESAGCQEWASVLTNPHLVELDITGNAL 839
QY 814 EDLGLRLCCOGLRHPVCRLLTLMKICRLTAACDELAETLSVNOQSLRELDLSNELGDL 873
DB 840 EDLGLRLCCOGLRHPVCRLLTLMKICRLTAACDELAETLSVNOQSLRELDLSNELGDL 899
QY 874 GVLLLCGLRHPTCKLQTLRLGICRLGSAACEGLSVVLQANHNRLRELDLSFNDLGDWGLW 933

```

```

; US-10-066-521-18
; Sequence 18, Application US/10066521
; Publication No. US20030027757A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Wang, Weiye
; APPLICANT: Blatcher, Maria
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN/NBS/LRR
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-334001
; CURRENT APPLICATION NUMBER: US/10/066,521
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/318,645
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/265,231
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1061
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-066-521-18

Query Match      99.3%; Score 5433; DB 14; Length 1061;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1034; Conservative 0; Mismatches 0; Indels 28; Gaps 2;

QY 1 MLRTAGRDGICRLSTYLELEAVELEKFKLYLGATATGEGKIPWGSMEKAGLEMAQLL 60
DB 1 MLRTAGRDGICRLSTYLELEAVELEKFKLYLGATATGEGKIPWGSMEKAGLEMAQLL 60
QY 61 ITHFGPEAWRLALSTFERINRKLWERGQREDLVR----- 96
DB 61 ITHFGPEAWRLALSTFERINRKLWERGQREDLVR----- 96
QY 97 ---DPOETRYDYVRKFRMLMEDNARLGEVCNLSHRYTRLLLVKHSNPMQVQQLDGTG 153
DB 121 PRKDPQSTYDYVRKFRMLMEDNARLGEVCNLSHRYTRLLLVKHSNPMQVQQLDGTG 180
QY 154 RGHARTVGHQASPIKIETLPEPDEERPEPPRTVVMQGAAGIGKSMIAHKVMDWDGKLF 213
DB 181 RGHARTVGHQASPIKIETLPEPDEERPEPPRTVVMQGAAGIGKSMIAHKVMDWDGKLF 240
QY 214 QGRFDYLYNCREMNSQATECSMDLI FSCWPEPSAPLOELIRVPERLLFIIDGDELK 273
DB 241 QGRFDYLYNCREMNSQATECSMDLI FSCWPEPSAPLOELIRVPERLLFIIDGDELK 300
QY 274 PSFHDPPQPCWCLCWEERKPTLELLNSLIRKLLPELSLITTRPTALEKHLRLEHPRHV 333
DB 301 PSFHDPPQPCWCLCWEERKPTLELLNSLIRKLLPELSLITTRPTALEKHLRLEHPRHV 360
QY 334 EILGFSAEARKEYFYKYFHNAEQAGQVFNVRDNEPLFTWCFVPLVCWVCTCLOOLEG 393
DB 361 EILGFSAEARKEYFYKYFHNAEQAGQVFNVRDNEPLFTWCFVPLVCWVCTCLOOLEG 420
QY 394 GGLLRQTSRTTAVYMLYLLSLMOPKPGAPLQPPNORGLCSLAADGLWNKILFEEQD 453
DB 421 GGLLRQTSRTTAVYMLYLLSLMOPKPGAPLQPPNORGLCSLAADGLWNKILFEEQD 480

```



```
QY 454 LRKHGLEDGVSFAFLNNWIFOKDINCERYYSFIHLSFOEFFAAMYIILDEGGAGPDD 513
DB 481 LRKHGLEDGVSFAFLNNWIFOKDINCERYYSFIHLSFOEFFAAMYIILDEGGAGPDD 540
QY 514 VTRLLTEYAFSERSFLATSRFLGFLNEETRSHLEKSLCWKSPHIIKMDLLOWTQSKAQ 573
DB 541 VTRLLTEYAFSERSFLATSRFLGFLNEETRSHLEKSLCWKSPHIIKMDLLOWTQSKAQ 600
QY 574 SDGSTLQOGSLEFFSCLYEIOEEFIQALSHFQIVVSNIASKMEHNVSSFCCLKRCSA 633
DB 601 SDGSTLQOGSLEFFSCLYEIOEEFIQALSHFQIVVSNIASKMEHNVSSFCCLKRCSA 660
QY 634 QVLHLYGATYSADGEDRARCAGAHNTLVQLRPVTLDDAYSEHLAAALCTNPNLIELS 693
DB 661 QVLHLYGATYSADGEDRARCAGAHNTLVQLRPVTLDDAYSEHLAAALCTNPNLIELS 719
QY 694 LYRNALSGRGVKKLCOGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 753
DB 720 LYRNALSGRGVKKLCOGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 779
QY 754 NGVFPFGMWLLCEGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 813
DB 780 NGVFPFGMWLLCEGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 839
QY 814 EDLGLRLCCOGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 873
DB 840 EDLGLRLCCOGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 899
QY 874 GVLLLCBGLRHPTCKLQTLRIGICRLGSAACEGLSVVLQANHNRELDLSFNDLGDWGLW 933
DB 900 GVLLLCBGLRHPTCKLQTLRIGICRLGSAACEGLSVVLQANHNRELDLSFNDLGDWGLW 959
QY 934 LLAEGLOHPACRLQKWLDSOGLTAKACENLYFTLGINQTLTDLYLTNNALGDTGVRLLC 993
DB 960 LLAEGLOHPACRLQKWLDSOGLTAKACENLYFTLGINQTLTDLYLTNNALGDTGVRLLC 1019
QY 994 KRLSHPGKLVLMFGMDLNKMTSHRLAALRVTKPYLDIGC 1035
DB 1020 KRLSHPGKLVLMFGMDLNKMTSHRLAALRVTKPYLDIGC 1061

RESULT 6
US-10-407-866-68
; Sequence 68, Application US/10407866
; Publication No. US20040002593A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; FILE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: 66654-10(IJ 5755)
; CURRENT APPLICATION NUMBER: US/10/407,866
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 60/370,538
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 1027
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-866-68

Query Match 96.1%; Score 5259.5; DB 15; Length 1027;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1000; Conservative 0; Mismatches 0; Indels 27; Gaps 1;

QY 1 MLRTAGRDGLCRSLTYLBELEAVALKKFKLYLGATLGECKIPWGSMEKAGPLEMAQLL 60
DB 1 MLRTAGRDGLCRSLTYLBELEAVALKKFKLYLGATLGECKIPWGSMEKAGPLEMAQLL 60
QY 61 ITHFGPEAWRLALSTFRINRKNLWERGQEDLVR----- 96
|||||
```

```
DB 61 ITHFGPEAWRLALSTFRINRKNLWERGQEDLVRDTPPGPSSLGNSQSTCLLEVSUVT 120
QY 97 ---DPQETRYDYVRKFRLEMDRNARLGEVCYNLSHRVTRLLLVKHSNPMVQOQLDGTG 153
DB 121 PRKDPQETRYDYVRKFRLEMDRNARLGEVCYNLSHRVTRLLLVKHSNPMVQOQLDGTG 180
QY 154 RGHARTVGHQASPIKIEITLPEPDERPEPRTVVMQGAAGIKGSMLAHKVMDLWADGKLF 213
DB 181 RGHARTVGHQASPIKIEITLPEPDERPEPRTVVMQGAAGIKGSMLAHKVMDLWADGKLF 240
QY 214 QGRFDYLFYINCREMNQSAECMODLIFSCWPPSPAPLOELIRVPERLLFIIDGFDELK 273
DB 241 QGRFDYLFYINCREMNQSAECMODLIFSCWPPSPAPLOELIRVPERLLFIIDGFDELK 300
QY 274 PSFHDPOGPWCLCWEKRPTELLNSLIRKLLPELSLITRPTALEKHLRLEHPRHV 333
DB 301 PSFHDPOGPWCLCWEKRPTELLNSLIRKLLPELSLITRPTALEKHLRLEHPRHV 360
QY 334 EILGFSBAERKEYFYKTFHNAEQAGVFNVDNEPLFTMCFVPLVVCWVCTCLOOQLEG 393
DB 361 EILGFSBAERKEYFYKTFHNAEQAGVFNVDNEPLFTMCFVPLVVCWVCTCLOOQLEG 420
QY 394 GGLLRQTSRTTAVYMLYLLSLMQPKCAPRLQPPNQGLCSLAADGLWNQKILFESQD 453
DB 421 GGLLRQTSRTTAVYMLYLLSLMQPKCAPRLQPPNQGLCSLAADGLWNQKILFESQD 480
QY 454 LRKHGLEDGVSFAFLNNWIFOKDINCERYYSFIHLSFOEFFAAMYIILDEGGAGPDD 513
DB 481 LRKHGLEDGVSFAFLNNWIFOKDINCERYYSFIHLSFOEFFAAMYIILDEGGAGPDD 540
QY 514 VTRLLTEYAFSERSFLATSRFLGFLNEETRSHLEKSLCWKSPHIIKMDLLOWTQSKAQ 573
DB 541 VTRLLTEYAFSERSFLATSRFLGFLNEETRSHLEKSLCWKSPHIIKMDLLOWTQSKAQ 600
QY 574 SDGSTLQOGSLEFFSCLYEIOEEFIQALSHFQIVVSNIASKMEHNVSSFCCLKRCSA 633
DB 601 SDGSTLQOGSLEFFSCLYEIOEEFIQALSHFQIVVSNIASKMEHNVSSFCCLKRCSA 660
QY 634 QVLHLYGATYSADGEDRARCAGAHNTLVQLRPVTLDDAYSEHLAAALCTNPNLIELS 693
DB 661 QVLHLYGATYSADGEDRARCAGAHNTLVQLRPVTLDDAYSEHLAAALCTNPNLIELS 720
QY 694 LYRNALSGRGVKKLCOGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 753
DB 721 LYRNALSGRGVKKLCOGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 780
QY 754 NGVFPFGMWLLCEGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 813
DB 781 NGVFPFGMWLLCEGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 840
QY 814 EDLGLRLCCOGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 873
DB 841 EDLGLRLCCOGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSG 900
QY 874 GVLLLCBGLRHPTCKLQTLRIGICRLGSAACEGLSVVLQANHNRELDLSFNDLGDWGLW 933
DB 901 GVLLLCBGLRHPTCKLQTLRIGICRLGSAACEGLSVVLQANHNRELDLSFNDLGDWGLW 960
QY 934 LLAEGLOHPACRLQKWLDSOGLTAKACENLYFTLGINQTLTDLYLTNNALGDTGVRLLC 993
DB 961 LLAEGLOHPACRLQKWLDSOGLTAKACENLYFTLGINQTLTDLYLTNNALGDTGVRLLC 1020
QY 994 KRLSHPG 1000
DB 1021 KRLSHPG 1027

RESULT 7
US-10-108-260A-3161
; Sequence 3161, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
```

| | | | |
|----|-----|---|------|
| Db | 840 | EDLGLRLLCQGLRHVPVCRRLTLWLKICRLTAACDELASTLSVNQSLRELDLSLNLGDL | 899 |
| Qy | 874 | GVLLICEGLRHPTCKLQTLRLGICRLGSAAECGLSVLQANHNLRELDLSFNDLGDWGLW | 933 |
| Db | 900 | GVLLICEGLRHPTCKLQTLR | 919 |
| Qy | 934 | LLAEGLOHPACRLQKLWLDSCGLTAKACENLYFTLGINQTLTDLTYLTNNALGDTGVRLLC | 993 |
| Db | 920 | -----LDSCGLTAKACENLYFTLGINQTLTDLTYLTNNALGDTGVRLLC | 962 |
| Qy | 994 | KRLSHPGCKLRVLWLFQMDLNKWTSHRLAALRVTKPYLDIGC | 1035 |
| Db | 963 | KRLSHPGCKLRVLWLFQMDLNKWTSHRLAALRVTKPYLDIGC | 1004 |

RESULT 8

US-10-407-866-90

; Sequence 90, Application US/10407866

; Publication No. US20040002593A1

; GENERAL INFORMATION:

; APPLICANT: Reed, John C.

; APPLICANT: Godzik, Adam

; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,

; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use

; FILE REFERENCE: 66654-10(LD 5755)

; CURRENT APPLICATION NUMBER: US/10/407,866

; CURRENT FILING DATE: 2003-04-04

; PRIOR APPLICATION NUMBER: US 60/370,538

; PRIOR FILING DATE: 2002-04-04

; NUMBER OF SEQ ID NOS: 129

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 90

; LENGTH: 635

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-407-866-90

Query Match 61.5%; Score 3364; DB 15; Length 635;

Best Local Similarity 100.0%; Pred. No. 6e-292;

Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0

| | | | |
|----|-----|---|-----|
| Qy | 1 | MLRTAGRDGLCRSLSTYLEEAEVLEAELKFKLYLTATATGELGKIPWGSMEKAGLEMAQLL | 60 |
| Db | 1 | MLRTAGRDGLCRSLSTYLEEAEVLEAELKFKLYLTATATGELGKIPWGSMEKAGLEMAQLL | 60 |
| Qy | 61 | ITHFGPEANRLALSTFERINRKDLWEGOREDLVRDPQSTYDYVRRKPRLMEDRNARL | 120 |
| Db | 61 | ITHFGPEANRLALSTFERINRKDLWEGOREDLVRDPQSTYDYVRRKPRLMEDRNARL | 120 |
| Qy | 121 | GEVNLSHRYTRILLVKEHSPMPQVQQLDTRGCHARTVGHQASPIKIETLPEDEERP | 180 |
| Db | 121 | GEVNLSHRYTRILLVKEHSPMPQVQQLDTRGCHARTVGHQASPIKIETLPEDEERP | 180 |
| Qy | 181 | EPPTVVMQGAAGIKGSLAHKVMLDWADGKLFQGRFDYLYINCRMNQASATCSMODL | 240 |
| Db | 181 | EPPTVVMQGAAGIKGSLAHKVMLDWADGKLFQGRFDYLYINCRMNQASATCSMODL | 240 |
| Qy | 241 | IFSCWPSPAPLOELIRVPERLLFIIDGFDLKPSPHDPPQWCLCWEKRPTTELLNSL | 300 |
| Db | 241 | IFSCWPSPAPLOELIRVPERLLFIIDGFDLKPSPHDPPQWCLCWEKRPTTELLNSL | 300 |
| Qy | 301 | IRKLLPELSLLITTRPTALEKHLRLEHPRHVEILGFSAEARKEYEYKVFHNAEQAGV | 360 |
| Db | 301 | IRKLLPELSLLITTRPTALEKHLRLEHPRHVEILGFSAEARKEYEYKVFHNAEQAGV | 360 |
| Qy | 361 | FNVTEDNEPLFTMCFVPLVCVWCTCLOQQLGEGGLLRQTSRTTAVYMLYLLSLMQPKP | 420 |
| Db | 361 | FNVTEDNEPLFTMCFVPLVCVWCTCLOQQLGEGGLLRQTSRTTAVYMLYLLSLMQPKP | 420 |
| Qy | 421 | GAPRLQPPNQRGLCSLAADGLMNQKILFEQDLRKHGLDGEDVSAFLNNIFOKDINCE | 480 |
| Db | 421 | GAPRLQPPNQRGLCSLAADGLMNQKILFEQDLRKHGLDGEDVSAFLNNIFOKDINCE | 480 |

| | | | |
|----|-----|---|------|
| Db | 840 | EDLGLRLLCQGLRHVPVCRRLTLWLKICRLTAACDELASTLSVNQSLRELDLSLNLGDL | 899 |
| Qy | 874 | GVLLICEGLRHPTCKLQTLRLGICRLGSAAECGLSVLQANHNLRELDLSFNDLGDWGLW | 933 |
| Db | 900 | GVLLICEGLRHPTCKLQTLR | 919 |
| Qy | 934 | LLAEGLOHPACRLQKLWLDSCGLTAKACENLYFTLGINQTLTDLTYLTNNALGDTGVRLLC | 993 |
| Db | 920 | -----LDSCGLTAKACENLYFTLGINQTLTDLTYLTNNALGDTGVRLLC | 962 |
| Qy | 994 | KRLSHPGCKLRVLWLFQMDLNKWTSHRLAALRVTKPYLDIGC | 1035 |
| Db | 963 | KRLSHPGCKLRVLWLFQMDLNKWTSHRLAALRVTKPYLDIGC | 1004 |

QY 481 RYYSFIHLSFOEFFAAMYIILDEGGAGPDQDVTRLLTEYAFSRSFLATSRFLGIL 540
DB 481 RYYSFIHLSFOEFFAAMYIILDEGGAGPDQDVTRLLTEYAFSRSFLATSRFLGIL 540
QY 541 NEETRSHLEKSLCWKVSPIHMKDILQWIOSKAQSDGSTLOQSLFFSCLYIOEEEFIQ 600
DB 541 NEETRSHLEKSLCWKVSPIHMKDILQWIOSKAQSDGSTLOQSLFFSCLYIOEEEFIQ 600
QY 601 QALSHFQVIVVSNIASKMEHVMVSFCLKRCRSAQV 635
DB 601 QALSHFQVIVVSNIASKMEHVMVSFCLKRCRSAQV 635

RESULT 9

US-09-965-621-64
; Sequence 64, Application US/09965621
; Publication No. US20030077699A1
; GENERAL INFORMATION:

; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowaki, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 603
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-965-621-64

Query Match 57.8%; Score 3163; DB 10; Length 603;
Best Local Similarity 100.0%; Pred. No. 5.7e-274;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 PNQGLCSLAADGLWNOKILFEEQDLRKHGIDGSDVSAFLNMNIFQKDINCERYYSFIHL 488
DB 1 PNQGLCSLAADGLWNOKILFEEQDLRKHGIDGSDVSAFLNMNIFQKDINCERYYSFIHL 60
QY 489 SFQEFFAAMYIILDEGGAGPDQDVTRLLTEYAFSRSFLATSRFLGILNEETRSHL 548
DB 61 SFQEFFAAMYIILDEGGAGPDQDVTRLLTEYAFSRSFLATSRFLGILNEETRSHL 120
QY 549 EKSICWKVSPIHMKDILQWIOSKAQSDGSTLOQSLFFSCLYIOEEEFIQALSHFQV 608
DB 121 EKSICWKVSPIHMKDILQWIOSKAQSDGSTLOQSLFFSCLYIOEEEFIQALSHFQV 180
QY 609 IVVSNIASKMEHVMVSFCLKRCRSAQVHLGYATYSADGEDRARCASAGAHLLVQLRPER 668
DB 181 IVVSNIASKMEHVMVSFCLKRCRSAQVHLGYATYSADGEDRARCASAGAHLLVQLRPER 240
QY 669 TVLLDAYSEHLAAALCTNPNIIELSLYRNALSGRGVKLLCOGLRHPNCKLQNLKRCRI 728
DB 241 TVLLDAYSEHLAAALCTNPNIIELSLYRNALSGRGVKLLCOGLRHPNCKLQNLKRCRI 300
QY 729 SSSACEDLSAALIANKNLTRMDLSGNGVGFPGMWLLCEGLRHPQCRQMIOLRKCOLES 788
DB 301 SSSACEDLSAALIANKNLTRMDLSGNGVGFPGMWLLCEGLRHPQCRQMIOLRKCOLES 360
QY 789 ACQEMASVLGTNPHLVELDITGNALDGLRLLCQGLRHPVCRRLTLWLKICRLTAACD 848
DB 361 ACQEMASVLGTNPHLVELDITGNALDGLRLLCQGLRHPVCRRLTLWLKICRLTAACD 420

QY 849 ELASTLSVNQSLRELDLSINELDGLVLLLCCEGLRHPCTKLTQTLRLGICRLGSAACEGLS 908
DB 421 ELASTLSVNQSLRELDLSINELDGLVLLLCCEGLRHPCTKLTQTLRLGICRLGSAACEGLS 480
QY 909 VVLQANHNRLDLSFNDLGDWGLWLAEGLOHPACRLQKWLDSOGLTAKACENLYFTL 968
DB 481 VVLQANHNRLDLSFNDLGDWGLWLAEGLOHPACRLQKWLDSOGLTAKACENLYFTL 540
QY 969 GINQTLTDLVLTNNALGDTGVRLLCRSLSHPSGCKRLVILWLFMDLNKMTSRLAALRYTK 1028
DB 541 GINQTLTDLVLTNNALGDTGVRLLCRSLSHPSGCKRLVILWLFMDLNKMTSRLAALRYTK 600
QY 1029 PYL 1031
DB 601 PYL 603

RESULT 10

US-10-407-866-64
; Sequence 64, Application US/10407866
; Publication No. US20040002593A1
; GENERAL INFORMATION:

; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; FILE REFERENCE: 66654-10(LJ 5755)
; CURRENT APPLICATION NUMBER: US/10/407,866
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 60/370,538
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 603
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-407-866-64

Query Match 57.8%; Score 3163; DB 15; Length 603;
Best Local Similarity 100.0%; Pred. No. 5.7e-274;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 PNQGLCSLAADGLWNOKILFEEQDLRKHGIDGSDVSAFLNMNIFQKDINCERYYSFIHL 488
DB 1 PNQGLCSLAADGLWNOKILFEEQDLRKHGIDGSDVSAFLNMNIFQKDINCERYYSFIHL 60
QY 489 SFQEFFAAMYIILDEGGAGPDQDVTRLLTEYAFSRSFLATSRFLGILNEETRSHL 548
DB 61 SFQEFFAAMYIILDEGGAGPDQDVTRLLTEYAFSRSFLATSRFLGILNEETRSHL 120
QY 549 EKSICWKVSPIHMKDILQWIOSKAQSDGSTLOQSLFFSCLYIOEEEFIQALSHFQV 608
DB 121 EKSICWKVSPIHMKDILQWIOSKAQSDGSTLOQSLFFSCLYIOEEEFIQALSHFQV 180
QY 609 IVVSNIASKMEHVMVSFCLKRCRSAQVHLGYATYSADGEDRARCASAGAHLLVQLRPER 668
DB 181 IVVSNIASKMEHVMVSFCLKRCRSAQVHLGYATYSADGEDRARCASAGAHLLVQLRPER 240
QY 669 TVLLDAYSEHLAAALCTNPNIIELSLYRNALSGRGVKLLCOGLRHPNCKLQNLKRCRI 728
DB 241 TVLLDAYSEHLAAALCTNPNIIELSLYRNALSGRGVKLLCOGLRHPNCKLQNLKRCRI 300
QY 729 SSSACEDLSAALIANKNLTRMDLSGNGVGFPGMWLLCEGLRHPQCRQMIOLRKCOLES 788
DB 301 SSSACEDLSAALIANKNLTRMDLSGNGVGFPGMWLLCEGLRHPQCRQMIOLRKCOLES 360
QY 789 ACQEMASVLGTNPHLVELDITGNALDGLRLLCQGLRHPVCRRLTLWLKICRLTAACD 848
DB 361 ACQEMASVLGTNPHLVELDITGNALDGLRLLCQGLRHPVCRRLTLWLKICRLTAACD 420
QY 849 ELASTLSVNQSLRELDLSINELDGLVLLLCCEGLRHPCTKLTQTLRLGICRLGSAACEGLS 908

Db 421 ELASTLSVNSQSLRELDLSNELGDLGVLLCEGLRHPTCKLQTLRLGICRLGSAACEGLS 480
Qy 909 VVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLWDSGGLTAKACENLYFTL 968
Db 481 VVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLWDSGGLTAKACENLYFTL 540
Qy 969 GINQTLTDLYLTNNALGDTGVRLLCKRLSHPGCKLRVLWLFQMDLNKMTSHSLAALRVTK 1028
Db 541 GINQTLTDLYLTNNALGDTGVRLLCKRLSHPGCKLRVLWLFQMDLNKMTSHSLAALRVTK 600
Qy 1029 PYL 1031
Db 601 PYL 603
RESULT 11
US-10-781-294-64
; Sequence 64, Application US/10781294
; Publication No. US20040142374A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/10/781,294
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US/09/965,621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 603
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-781-294-64
Query March 57.8%; Score 3163; DB 16; Length 603;
Best Local Similarity 100.0%; Pred. No. 5.7e-274;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 429 PNQGLCSLAADGLWNNKILFEQDLRKHGLDGEDVSAFLNNIFQKQINCERYYSFIHL 488
Db 1 PNQGLCSLAADGLWNNKILFEQDLRKHGLDGEDVSAFLNNIFQKQINCERYYSFIHL 60
Qy 489 SFQEFFAAMYILDEGGGAGPDQVTRLITVAFSERSFLALTSRFLGLNEETRSHL 548
Db 61 SFQEFFAAMYILDEGGGAGPDQVTRLITVAFSERSFLALTSRFLGLNEETRSHL 120
Qy 549 EKSLCWKVSPIH KMDLQWIOSKAQSDGSTLQGSLEFFSCLYEIOEERFIOQALSHFOV 608
Db 121 EKSLCWKVSPIH KMDLQWIOSKAQSDGSTLQGSLEFFSCLYEIOEERFIOQALSHFOV 180
Qy 609 IWSNIASKMEHMYVSFLCKRCRAQVHLVYGATYSADGEDRARCAGAHLLVQLRPER 668
Db 181 IWSNIASKMEHMYVSFLCKRCRAQVHLVYGATYSADGEDRARCAGAHLLVQLRPER 240
Qy 669 TVLLDAYSEHLAAALCTNPNIIEISLYRNAIGSGVKLLCGGLRHPTCKLQTLRLGICRLGSAACEGLS 728
Db 241 TVLLDAYSEHLAAALCTNPNIIEISLYRNAIGSGVKLLCGGLRHPTCKLQTLRLGICRLGSAACEGLS 300
Qy 729 SSSACEDLSAALIANKNLTRMDLSGNGVGFPGWMLLCEGLRHPTCKLQTLRLGICRLGSAACEGLS 788
Db 301 SSSACEDLSAALIANKNLTRMDLSGNGVGFPGWMLLCEGLRHPTCKLQTLRLGICRLGSAACEGLS 360
Qy 789 ACOEMASVLGTNPVPHLVELDITGNALEDLGLRLLCGGLRHPTCKLQTLRLGICRLGSAACEGLS 848

Db 361 ACOEMASVLGTNPVPHLVELDITGNALEDLGLRLLCGGLRHPTCKLQTLRLGICRLGSAACEGLS 420
Qy 849 ELASTLSVNSQSLRELDLSNELGDLGVLLCEGLRHPTCKLQTLRLGICRLGSAACEGLS 908
Db 421 ELASTLSVNSQSLRELDLSNELGDLGVLLCEGLRHPTCKLQTLRLGICRLGSAACEGLS 480
Qy 909 VVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLWDSGGLTAKACENLYFTL 968
Db 481 VVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLWDSGGLTAKACENLYFTL 540
Qy 969 GINQTLTDLYLTNNALGDTGVRLLCKRLSHPGCKLRVLWLFQMDLNKMTSHSLAALRVTK 1028
Db 541 GINQTLTDLYLTNNALGDTGVRLLCKRLSHPGCKLRVLWLFQMDLNKMTSHSLAALRVTK 600
Qy 1029 PYL 1031
Db 601 PYL 603
RESULT 12
US-10-162-335-22
; Sequence 22, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Hjalt, Tord
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zernhusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Metho
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10/162,335
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883

```
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM:
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 22
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-335-22

Query Match          54.9%; Score 3005; DB 15; Length 582;
Best Local Similarity 99.3%; Pred. No. 7,8e-260;
Matches 578; Conservative 0; Mismatches 0; Indels 4; Gaps 3;

Qy 373 MCFPLVWVWVCTCLOQLEGGLRQTSRTTAVYMLYLSLMPKPGAPRLQPPNQ 432
Db 1 MCFPLVWVWVCTCLOQLEGGLRQTSRTTAVYMLYLSLMPKPGAPRLQPPNQ 60
Qy 433 GLCSLAADGLNOKILFEQDLRKHGLDGEDVSFAFLNNIFQKDCINCYRYSFHLSFOE 492
Db 61 GLCSLAADGLNOKILFEQDLRKHGLDGEDVSFAFLNNIFQKDCINCYRYSFHLSFOE 120
Qy 493 FFAAMYIILDEGEAGAGPDQVTRLLTEYAFSERSFLALTSRFLFGLLNEETRSHLEKSL 552
Db 121 FFAAMYIILDEGEAGAGPDQVTRLLTEYAFSERSFLALTSRFLFGLLNEETRSHLEKSL 180
Qy 553 CWKVSPIHKMDLLQWIOSKAQSDGSTLQGSLEFPSCLYEIOEBSFIOQALSHPFQVIVS 612
Db 181 CWKVSPIHKMDLLQWIOSKAQSDGSTLQGSLEFPSCLYEIOEBSFIOQALSHPFQVIVS 240
Qy 613 NIASKMEHWSFCLKRCRSQVLLHYGATYSADGEDRARSAGAHILLVOLRPERTVILL 672
Db 241 NIASKMEHWSFCLKRCRSQVLLHYGATYSADGEDRARSAGAHILLVOLRPERTVILL 300
Qy 673 DAYSEHLAAALCTPNLIELSLYRNALSGRVKLLCOGLRHPNCKLQNL-ELKRCRISSS 731
Db 301 DAYSEHLAAALCTPNLIELSLYRNALSGRVKLLCOGLRHPNCKLQNL-ELKRCRISSS 360
Qy 732 ACEDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPQCRLOMIQLRKCQLESAGCQ 791
Db 361 ACEDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPQCRLOMIQLRKCQLESAGCQ 420
Qy 792 EMASVLGTPNPHVELDITGNALEDGLRLICQGLRHPVCRRLTLW-LKICRLTAACADE 849
Db 421 EMASVLGTPNPHVELDITGNALEDGLRLICQGLRHPVCRRLTLW-LKICRLTAACADE 480
Qy 850 LASTLSVNSQSLRELDLSNELGDLGVLLLCGLRHPTCKLQTL-RLGICRLGSAACEGLS 908
Db 481 LASTLSVNSQSLRELDLSNELGDLGVLLLCGLRHPTCKLQTL-RLGICRLGSAACEGLS 540
Qy 909 VVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKLW 950
Db 541 VVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKLW 582

RESULT 13
US-10-467-397-9
; Sequence 9, Application US/10467397
; Publication No. US20040137448A1
; GENERAL INFORMATION:
; APPLICANT: THORNTON, Michael; HAFALIA, April J.A.;
; APPLICANT: LU, Dyung Aina M.; ARVIZU, Chandra S.;
; APPLICANT: SWARNAKAR, Anita; LU, Yan;
; APPLICANT: WARREN, Bridget A.; BAUGHN, Mariah R.;
; APPLICANT: TANG, Y. Tom; LEE, Ernestine A.;
; APPLICANT: YAO, Monique G.; RAMKUMAR, Jayalaxmi;
; APPLICANT: KHAN, Farrah A.; GANDHI, Ameena R.;
; APPLICANT: DING, Li; YUE, Henry;
; APPLICANT: GIETZEN, Kimberly J.; CHAWLA, Narinder K.;
; APPLICANT: THANGAVELU, Kavitha; ELLIOTT, Vicki S.
; TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS
; FILE REFERENCE: PI-0368 USN
; CURRENT APPLICATION NUMBER: US/10/467,397
```

```
; CURRENT FILING DATE: 2003-08-06
; PRIOR APPLICATION NUMBER: PCT/US02/03844
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/268,118
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/270,963
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,858
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/271,194
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/274,071
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/283,496
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 60/344,650
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 429930CD1
US-10-467-397-9

Query Match          54.0%; Score 2953.5; DB 16; Length 565;
Best Local Similarity 99.8%; Pred. No. 3.1e-255;
Matches 565; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 470 MNIFQKDCINCYRYSFHLSFOEFAAMYIILDEGEAGAGPDQVTRLLTEYAFSERSFL 529
Db 1 MNIFQKDCINCYRYSFHLSFOEFAAMYIILDEGEAGAGPDQVTRLLTEYAFSERSFL 60
Qy 530 ALTSRFLFGLLNEETRSHLEKSLCWKVSPIHKMDLLQWIOSKAQSDGSTLQGSLEFPSC 589
Db 61 ALTSRFLFGLLNEETRSHLEKSLCWKVSPIHKMDLLQWIOSKAQSDGSTLQGSLEFPSC 120
Qy 590 LYEQEBSFIOQALSHPFQVIVSNIAKMEHWSFCLKRCRSQVLLHYGATYSADGED 649
Db 121 LYEQEBSFIOQALSHPFQVIVSNIAKMEHWSFCLKRCRSQVLLHYGATYSADGED 180
Qy 650 RARCSAGAHILLVOLRPERTVLLDAYSEHLAAALCTPNLIELSLYRNALSGRVKLLCQ 709
Db 181 RARCSAGAHILLVOL-PERTVLLDAYSEHLAAALCTPNLIELSLYRNALSGRVKLLCQ 239
Qy 710 GLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLR 769
Db 240 GLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLR 299
Qy 770 HPQCRLOMIQLRKCQLESAGCQEMASVLGTPNPHVELDITGNALEDGLRLICQGLRHPV 829
Db 300 HPQCRLOMIQLRKCQLESAGCQEMASVLGTPNPHVELDITGNALEDGLRLICQGLRHPV 359
Qy 830 CRLRTLMLKICRLTAACADELASTLSVNSQSLRELDLSNELGDLGVLLLCGLRHPTCKL 889
Db 360 CRLRTLMLKICRLTAACADELASTLSVNSQSLRELDLSNELGDLGVLLLCGLRHPTCKL 419
Qy 890 QTLRLGICRLGSAACEGLSVVVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKL 949
Db 420 QTLRLGICRLGSAACEGLSVVVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKL 479
Qy 950 WLDSGGLTAKACENLYFTLGINOTLTDLYLTNNALGDTGVLLCKRSLHPCCKLRVLWLF 1009
Db 480 WLDSGGLTAKACENLYFTLGINOTLTDLYLTNNALGDTGVLLCKRSLHPCCKLRVLWLF 539
Qy 1010 GMDLNKMTSHSLAALRVTKPYLDIGC 1035
Db 540 GMDLNKMTSHSLAALRVTKPYLDIGC 565
```

```
RESULT 14
US-10-162-335-24
; Sequence 24, Application US/10162335
; Publication No. US20040009480A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Hjal, Tord
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Macdougall, John R.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Stone, David J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zerkusen, Bryan D.
; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Method
; FILE REFERENCE: 21402-377 B
; CURRENT APPLICATION NUMBER: US/10/162,335
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; PRIOR FILING DATE: 2001-06-26
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 201
; SEQ ID NO 24
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-335-24

Query Match 49.2%; Score 2694.5; DB 15; Length 521;
Best Local Similarity 90.1%; Pred. No. 4.3e-232;
Matches 521; Conservative 0; Mismatches 0; Indels 57; Gaps 1;

Qy 373 MCFVPLVWVCTCLOQLEGGLRQTSRTTAVTMYLLSLMOPKGPAPLQPPNQR 432
Db 1 MCFVPLVWVCTCLOQLEGGLRQTSRTTAVTMYLLSLMOPKGPAPLQPPNQR 60

Qy 433 GLCSLAADGLWNOKILFEEDOLRKHGJGDGVSAFLNNMIFQDINCERYYSFIHLSFOE 492
Db 61 GLCSLAADGLWNOKILFEEDOLRKHGJGDGVSAFLNNMIFQDINCERYYSFIHLSFOE 120
```

```
Qy 493 FFAAMYILDEGEGAGPDQDVTRLLTEYAFSRSFLALTSRFLFGLNNEETSHLEKSL 552
Db 121 FFAAMYILDEGEGAGPDQDVTRLLTEYAFSRSFLALTSRFLFGLNNEETSHLEKSL 180

Qy 553 CWKVSPIKMDLLQWIOSKQSDGSTLQOQSLFFSCFYIEIEEFIOQALSHFQVIVVS 612
Db 181 CWKVSPIKMDLLQWIOSKQSDGSTLQOQSLFFSCFYIEIEEFIOQALSHFQVIVVS 240

Qy 613 NIASKMEHVSFCLKECRSAQVILHYGATYSADGEDRARCSCAGAHLLVQLRPERTVLL 672
Db 241 NIASKMEHVSFCLKECRSAQVILHYGATYSADGEDRARCSCAGAHLLVQLRPERTVLL 300

Qy 673 DAYSEHLAAALCTNPNIIEISLYENALGSRGVKLLCQGLRHPNCKLQNLKRCRISSSA 732
Db 301 DAYSEHLAAALCTNPNIIEISLYENALGSRGVKLLCQGLRHPNCKLQNLKRCRISSSA 360

Qy 733 CEDLSAALIANKNLTTRMDLSNGVGFPGMMLLCEGLRHPQCRLOMIQLRKQOLSSGACQE 792
Db 361 CEDLSAALIANKNLTTRMDLSNGVGFPGMMLLCEGLRHPQCRLOMIQLRKQOLSSGACQE 407

Qy 793 MASVLGTNPHLVELDLTGNALEDLGLRLCQGLRHPVCRLRTLMLKICRLTAACDELAS 852
Db 408 MASVLGTNPHLVELDLTGNALEDLGLRLCQGLRHPVCRLRTLMLKICRLTAACDELAS 423

Qy 853 TILSVNQSLRELDLSLNELDGLVLLCEGLRHPCKLOTLRIGTCRIGSAACEGLSVVLQ 912
Db 424 TILSVNQSLRELDLSLNELDGLVLLCEGLRHPCKLOTLRIGTCRIGSAACEGLSVVLQ 483

Qy 913 ANHNRLRELDLSFNDIGDWGLWLLAEGLOHPACRLQKLW 950
Db 484 ANHNRLRELDLSFNDIGDWGLWLLAEGLOHPACRLQKLW 521

RESULT 15
US-10-127-516-5
; Sequence 5, Application US/10127516
; Publication No. US20020187922A1
; GENERAL INFORMATION:
; APPLICANT: Manji, Gulam A.
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE PYRIN DOMAIN
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-370001
; CURRENT APPLICATION NUMBER: US/10/127,516
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 10/027,629
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 09/964,955
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 09/653,901
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/506, 067
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-127-516-5

Query Match 43.8%; Score 2396; DB 13; Length 1034;
Best Local Similarity 46.1%; Pred. No. 7e-205;
Matches 495; Conservative 175; Mismatches 303; Indels 100; Gaps 10;

Qy 11 CRLSTYLELEAVEVLEKFKLYL-GTATELGEKIPMGSMKAGFLEMAQLLIHFGPEEA 69
Db 6 CKLARYLEDLEDVLDKFKHLEDYPPQKICPLPRQTEKADHVDIATLMIDFNGBEKA 65

Qy 70 WRLALSTFERINRKLWERCQED- 93
Db 66 WAMAVWTFIAAIIINRDLRYEKAKRDEPKWGNARVSNPTVICOEDSIEEWNGLLRSRI 125
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 14:44:08 ; Search time 50 Seconds
(without alignments)
1991.687 Million cell updates/sec

Title: US-10-781-294-24
Perfect score: 5472
Sequence: 1 MLRTAGRDGLCRSLSTYLEEL.....MTHSRLAALRVTKPYLDIGC 1035

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|----------|--------------------|
| 1 | 1314.5 | 24.0 | 1192 | 2 T17255 | hypothetical prote |
| 2 | 1160.5 | 21.2 | 1111 | 2 A59000 | mater protein (imp |
| 3 | 724 | 13.2 | 461 | 2 A31858 | ribonuclease-angio |
| 4 | 717 | 13.1 | 456 | 2 S20597 | ribonuclease inhib |
| 5 | 709 | 13.0 | 456 | 2 A31857 | ribonuclease inhib |
| 6 | 494.5 | 9.0 | 483 | 2 S27880 | Nasopressin recept |
| 7 | 377.5 | 6.9 | 1130 | 2 A48843 | MHC class II trans |
| 8 | 258 | 4.7 | 1004 | 2 T31665 | hypothetical prote |
| 9 | 216.5 | 4.0 | 1075 | 2 T31668 | hypothetical prote |
| 10 | 183.5 | 3.4 | 1232 | 2 A54748 | neuronal apoptosis |
| 11 | 180 | 3.3 | 312 | 2 B97746 | hypothetical prote |
| 12 | 179.5 | 3.3 | 545 | 2 T52068 | RAN GTPase-activat |
| 13 | 178 | 3.3 | 533 | 2 T52063 | RAN GTPase-activat |
| 14 | 177.5 | 3.2 | 506 | 2 A45841 | T-complex-associat |
| 15 | 172.5 | 3.2 | 618 | 2 T48193 | hypothetical prote |
| 16 | 171 | 3.1 | 1447 | 2 T42628 | neuronal apoptosis |
| 17 | 170 | 3.1 | 568 | 2 F86291 | hypothetical prote |
| 18 | 169.5 | 3.1 | 789 | 2 T52067 | hypothetical prote |
| 19 | 169.5 | 3.1 | 1121 | 2 T02764 | myosin-I binding p |
| 20 | 162 | 3.0 | 631 | 2 C89243 | protein F28C1.3 (i |
| 21 | 162 | 3.0 | 631 | 2 T21471 | hypothetical prote |
| 22 | 161.5 | 3.0 | 589 | 2 A36983 | RNAI homolog fugi |
| 23 | 160 | 2.9 | 998 | 2 T23427 | hypothetical prote |
| 24 | 158 | 2.9 | 535 | 2 T48102 | RAN GTPase activat |
| 25 | 156.5 | 2.9 | 589 | 2 T52070 | RNAI protein homol |
| 26 | 155 | 2.8 | 587 | 2 J05300 | RAN GTPase activat |
| 27 | 152.5 | 2.8 | 2493 | 2 A55481 | adenylate cyclase |
| 28 | 152 | 2.8 | 1010 | 2 T36383 | probable large ATP |
| 29 | 142.5 | 2.6 | 1253 | 2 T45787 | disease resistance |

| | | | | | |
|----|-------|-----|------|----------|--------------------|
| 30 | 142 | 2.6 | 934 | 2 T05201 | hypothetical prote |
| 31 | 140.5 | 2.6 | 1389 | 2 T13852 | gene wheeler prote |
| 32 | 140 | 2.6 | 1039 | 2 A85096 | hypothetical prote |
| 33 | 139.5 | 2.5 | 271 | 2 D84586 | hypothetical prote |
| 34 | 138 | 2.5 | 526 | 2 C84552 | hypothetical prote |
| 35 | 137.5 | 2.5 | 1237 | 2 A81915 | hypothetical prote |
| 36 | 137.5 | 2.5 | 1385 | 2 T13887 | tir protein - frui |
| 37 | 136.5 | 2.5 | 1286 | 2 T33476 | hypothetical prote |
| 38 | 135.5 | 2.5 | 1066 | 2 T15864 | hypothetical prote |
| 39 | 135 | 2.5 | 592 | 2 T52139 | LRR-containing F-b |
| 40 | 135 | 2.5 | 1217 | 2 T52348 | disease resistance |
| 41 | 134.5 | 2.5 | 906 | 2 G96621 | probable disease r |
| 42 | 134.5 | 2.5 | 957 | 2 E84547 | probable disease r |
| 43 | 134 | 2.4 | 607 | 2 F96598 | protein F20N2.2 (i |
| 44 | 133 | 2.4 | 1110 | 2 F84547 | probable disease r |
| 45 | 132.5 | 2.4 | 907 | 2 J0176 | orphan G protein-c |

ALIGNMENTS

RESULT 1

T17255
hypothetical protein DKFZp586O1822.1 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T17255
R:Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18722
A:Accession: T17255
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1192 <KOE>
A:Cross-references: UNIPROT:Q9C000; EMBL:AL117470
A:Experimental source: adult uterus; clone DKFZp586O1822
C:Genetics:
A>Note: DKFZp586O1822.1

Query Match 24.0% Score 1314.5; DB 2; Length 1192;
Best Local Similarity 39.9%; Pred. No. 2.3e-93;
Matches 323; Conservative 119; Mismatches 293; Indels 75; Gaps 19;

| | | | |
|----|-----|--|-----|
| Qy | 133 | LLLVKEH---SNPMQVQ--QLDTRGHARTVGHQASPIKTIETLFPDEPBPPTTVV | 187 |
| Db | 1 | LLLQRPSPQDPLVKESWPDYVEENRGL-----IEIRDLFGPLDTQF-PRIV | 50 |
| Qy | 188 | MQGAAGIKGMLAHKVMLDWADKLFQGRFDYLPYINCREMNGSATECSMODLIFSCWPE | 247 |
| Db | 51 | LQGAAGIKGKTLARQVKEAMGRGQLYGRFQHVYFSCRELAQSKV-VSLAELIGKDGTA | 109 |
| Qy | 248 | PSAPLOELIRVPERLLFIIDGDELKPSPHDPQGPWGLCWEEKPTLLNLSLRKKLLP | 307 |
| Db | 110 | TPAPIRILSRPERLLFIIDGDELKPSPHDPQGPWGLCWEEKPTLLNLSLRKKLLP | 169 |
| Qy | 308 | ELSLIITRTALEKHLRLHPRHVEILGFSAERKEYFYKYFHNAEQAGQVNVYRDN | 367 |
| Db | 170 | EASFLITARTALQNLPSELEQARWVEVLGFSSSRKEYFYRYTDERQAIRAFLVKSN | 229 |
| Qy | 368 | EPLFTMCFVPLVCWVVTCTLOQLEGGLLRQTSRTTAVMYLLSLMQPKGAPRIQP | 427 |
| Db | 230 | KELWALCLVPWSWLAFTCLMQMKRKEKLTLSKTTTTLCHVLAQALQAPLGPQL-- | 287 |
| Qy | 428 | PPNORIGCSLAADGLNNQKILFEEQDLRKHGLDGEDVSATFNWNIQKDCINERYYSFIH | 487 |
| Db | 288 | -----RDLCSSAAEGIMQKKTLPSPDDLKHKGLDGAIIITFLKMGILQEH-PIPIUSYFIH | 342 |
| Qy | 488 | LSFOEFPAAMYIILDEGEGGAGPDQ-----DVTLLIETAYFSPERSFLATLSRFLFLGLNEE | 543 |
| Db | 343 | LCQEFPAAMSYVL-EDEKGRGKHSNCIIIDEKLTLEYAGI-HGLFGASTTFFLLGLLSDE | 400 |
| Qy | 544 | TRSHLEKSLCWKVSPIHKMDLLOWIQSKAQSDGSLTQOGSLEFPSCFYIETQEEFFIOAAL | 603 |

401 GEREMENIFHCRLSQ--GRNLQWVPSLQ-----LILQPHSLSLHCLYETRNKTLFTQVM 454
Db
604 SHFOVIVSVNIAKMEHMYSSFCCKRCSAQVHLHYGATYSADGEDRARCAGATHLLVQ 663
Qy
455 AHEEMGMC-VETDMELLVCTFCIKPSRHVKQLIT-----BGRQHRSTWSTWVVL 505
Db
664 LRPBTVLLDAYSEHLAAALCTNPNIELSLYRNALSGVKGKLLCOGLRHPNCKLQNLRL 723
Qy
506 FR--WVPTVTDAYWQILFVSLKVTNKLKELDLGSLNSLHSAVKSCLKTLPKCLLETLR 563
Db
724 KRCRISASACDLASALIANKNLTRMDLSCNGVGFPGMMLLCEGLRHPNCKLQNLRL 783
Qy
564 AGCGHTADCKDLAFGRANQTLTDLSPNVLTDAKXHLQORLQPSCKLQRLQVSC 623
Db
784 QLESGACQEMASVLGTNPVHVELDLTGNVALEDGLRLCOGLRHPNCKLQNLRL 843
Qy
624 GLTSDCCDLASVLSASPSLKLQNLDDVGVRLLCGLRHPNCKLQNLRL 677
Db
844 AAACDELASTLSVNSQSLRELDLSNEL-----GDLGVLLLCCEGL-----RHPCKLQTLR 893
Qy
678 ----DQTLSDQMQLALBOEKPOLLIIFSRKPSVMTPTTEGLDTGEMSNSTSLKQRL 733
Db
894 LGICRLGSAACEGLSVVLOANHNRLRELDLS 923
Qy
734 LGSERAAHV-----AQNKLKLDVS 754
Db
RESULT 2
A59000
mater protein [imported] - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
R;Tong, Z.B.; Nelson, L.M.
Endocrinology 140, 3720-3726, 1999
A:Title: A mouse gene encoding an oocyte antigen associated with autoimmune premature ovulation
A:Reference number: A59000; MUID:99360614; PMID:10433232
A:Accession: A59000
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1111 <KUR>
A:Cross-references: UNIPROT:Q9RLM5; GB:AF074018; NID:G5802697; PIDN:AA051762.1; PID:G5802697
C:Genetics:
A:Gene: Mater
Query Match 21.2%; Score 1160.5; DB 2; Length 1111;
Best Local Similarity 31.4%; Pred. No. 1.9e-81;
Matches 293; Conservative 165; Mismatches 407; Indels 69; Gaps 13;
Qy 162 HQASP-IK-IETLEPDEEPERPTVWQGAAGICKSMIAKVMLDWADGKLFQGRFDY 219
Db 168 HYDSPMKLLSDAFKP-YQKTFQPHITLHGRFGVKSALARSIVLWAGKLFQ-KMSF 225
Qy 220 LFYINCRMNOSATECSMDLIFSCWPEPSAPLOELIRVPERLLFIIDGDELKPSFHPD 279
Db 226 VIFPSVREIKWT-EKSSLAQLIAKECDSDMLVTKIMSQPERLLFVIDGLDMDSDVLQHD 284
Qy 280 QCPMCLCWEKRPTELLNSLRKILPELSLITTRPTALEKHLRLHPRHVEILGFS 339
Db 285 DMTLSRDNDKQPIYILMSLRKALLPQSFLLITTRNTGLEKLSKMSVSPYLVEGLS 344
Qy 340 EAERKEYFYKFNHAEQAGQVFNVRDNEPLFTWCFVPLVCVVVCTCLOQLBEGGGLRQ 399
Db 345 AGRRQLVLENISNDRIOVFHSLIENHQLFDQCAPSVCSLVCEALQQLKGRCTL 404
Qy 400 TGRITTVAV---MLYLLSIMPQKCAPLQPPNQGRGLCSLAADGLNOKILPEEODLRK 456
Db 405 PCOTLTGLYATLVFHLTKRPSQALSQEBQITLVGLCMWAGVMTWMSVFDLDDKN 464
Qy 457 HGLDGDGVAFTNMI-FOKDINCERYYSFIHLSFQEFFAAMYIYLBEGGGAGPDQ--- 512
Db 465 YLSKESEIALFHMILLOVGHNSBQCYVFSHLSLDQFFAALYVYL---EGLEEWNHQFC 521

513 --DVTRLLTEYAFSERSFLATSRFLCGLLNEETRSHELSKLCWKVSPHIKMDLLQMIQS 570
Db 522 FIENQRSIMEVTRTDDTLGKRFGLMKDKILKTLEVLFEYVPTVBOQLQHWVSL 581
Qy 571 KQSDGSTLQOQSDEFFSCYIEQIEEFIOALSHQFVIVVSNIAKMEHMYSSFCCKR 630
Db 582 IAQVNGTSPMDTDAFYCLPESODEEFVGGALKRFOEVWLL-INQRMDLKVSSYCLKHC 640
Qy 631 RSAQVHL-----YGATYSADGE----- 648
Db 641 QNLKAIKVRDIRDLSVDNTLELCPVTVTOCKPPLMEWGNFCVLSGLSLNKLKELDLG 700
Qy 649 DRARCSAGATHLLVQLRPE-----RTVLDAYSEHLAAALCTNPNIELSLYRNAL 699
Db 701 DSILSQPANKILCLELRNQCRIQKLTFRSAEVSGLKHLWKLFSNQNLKYLNLGNTPM 760
Qy 700 GSRGVKLLCOGLRHPNCKLQNLRLKRCRISASACDLASALIANKNLTRMDLSCNGVGF 759
Db 761 KDDMKLACEALKHPKCSVETRLDSCELTIIGYEMISTLLISTTLTKCLSLAKNRVGVK 820
Qy 760 GMLLCEGLRHPNCKLQNLRLKRCRISASACDLASALIANKNLTRMDLSCNGVGF 819
Db 821 SMISLGNALSSMCLLQKLLDNCGLTPASCHLLVSALFSNQNLTHLCLSNNSLGTGVQ 880
Qy 820 LLCQGLRHPNCKLQNLRLKRCRISASACDLASALIANKNLTRMDLSCNGVGF 879
Db 881 QLCQFLNPECALQRLILNHCNIIVDDAYGFLAMRLANNNTKLTSLTMTVPVGDGAMKLLC 940
Qy 880 EGLRHPNCKLQNLRLKRCRISASACDLASALIANKNLTRMDLSCNGVGF 939
Db 941 EALKEPTCYLQLELVDQQLTQNCEDLACMIITTKHLKSLDGLNNALGKGVITLCEGL 1000
Qy 940 QHPACRLCKLWLDSCGLTAKACENLYTGLINQTLDTLYLTNNALGDTGVRLCKLSHP 999
Db 1001 KQSSSLRLRLGLGACKLTSCNCEALSALSCNPHLSNLVKNDFSTSGMLKLSAFQCP 1060
Qy 1000 GCKRLVWLFGNDLKNMTHSLAALRVTKPYLDI 1033
Db 1061 VSNLGIIGLWQOEYVARVRQRLEVEFVKPHVVI 1094
RESULT 3
A31858
ribonuclease-angiogenin inhibitor - human
N;Alternate names: ribonuclease inhibitor, placental
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C:Accession: A31858; S02012; S23933; S48636; T47188
R;Lee, F.S.; Fox, E.A.; Zhou, H.M.; Strydom, D.J.; Vallee, B.L.
Biochemistry 27, 8545-8553, 1988
A:Title: Primary structure of human placental ribonuclease inhibitor.
A:Reference number: A31858; MUID:89118269; PMID:3219362
A:Accession: A31858
A:Molecule type: mRNA
A:Residues: 1-461 <LEE>
A:Cross-references: UNIPROT:P13489; GB:M22414; NID:G186260; PIDN:AAA59130.1; PID:G30704.
R;Schneider, R.; Schneider-Scherzer, E.; Thurnher, M.; Auer, B.; Schweizer, M.
EMBO J. 7, 4151-4156, 1988
A:Title: The primary structure of human ribonuclease/angiogenin inhibitor (RAI) discloses
A:Reference number: S02012; MUID:89210799; PMID:3243277
A:Accession: S02012
A:Molecule type: mRNA
A:Residues: 1-422, 'SE', 425-461 <SCH>
A:Cross-references: EMBL:X13973; NID:G35843; PIDN:CAA32151.1; PID:G35844
A>Note: part of this sequence, including the carboxyl end of the mature protein, was co-
R;Crevel-Thieffry, I.; Cotterill, S.; Schuller, E.
Biochim. Biophys. Acta 1122, 107-112, 1992
A:Title: Characterisation of a cryptic peptide from human placental ribonuclease inhibi-
A:Reference number: S23933; MUID:92338217; PMID:1633192
A:Accession: S23933
A:Molecule type: protein
A:Residues: 174-195 <CRE>
R;Nadano, D.; Yasuda, T.; Takeshita, H.; Uchida, K.; Kishi, K.

| | Query Match | 13.0% | Score 709; | DB 2; | Length 456; |
|----|-----------------------|--------------------|--------------------|------------------------|--------------------|
| | Best Local Similarity | 42.7%; | Pred. No. 5.8e-47; | | |
| | Matches 151; | Conservative 60; | Mismatches 143; | Indels 0; | Gaps 0; |
| Qy | 671 | LLDAYSEHAAALCTNPNI | LSIYRNALSGRGVKLL | COGLRHFNCKLQNLRLKRCRIS | 730 |
| | | : : : : | : : : : | : : : : | : : : : |
| Db | 35 | LTEHCXDI | GSALFRANPSLT | TELCRTNELGDAGVHLV | LGQSPCKIQLSLQNC |
| | | : : : : | : : : : | : : : : | : : : : |
| Qy | 731 | SACEDLSAAALIANKNL | TRMDLSGNGVFFP | GMLLCEGLRHFPQCR | LOMIQLRKCOLES |
| | | : : : : | : : : : | : : : : | : : : : |
| Db | 95 | AGCOVLPT | LRSPPTLRHL | SLNDPLGDAGURL | LLCEGLDPOCHLEK |
| | | : : : : | : : : : | : : : : | : : : : |
| Qy | 791 | QEWASVLGTNP | PHLVLEPLDTGNALED | LGRLLCQGLRHFPVCR | LRLTTLWIKICRLTAAAC |
| | | : : : : | : : : : | : : : : | : : : : |

Wed Feb 9 10:22:15 2005

Db 155 EPLASVLRATRALKELTVSNNDIGEAGARVLGGIADLSACOLETLELNCGLTPANCKDL 214
Qy 851 ASTLSVNQSLRELDLSNELGDLGVLLCEGLRHPCKLQTLRLGICRLGSAACEGLSVV 910
Db 215 CGIVASQASURELDLGSNGLDGAGIAELCPGLSPASRLKTLWLWECDDTAGCRDLCEV 274
Qy 911 LQAHNHLRELDLGFNDLGDWGLWLLAEGLOHPACRIQLKWLDSGCGITAKACENLYFTLGI 970
Db 275 LQAKETLKSLAGNKLGDGELLCELLQPCQLESLEWVKSCSLTAACQHVSLMLTQ 334
Qy 971 NOTLTDLYLNALGDTGVRLCKRLSHPCCKLRVLWLFMDLNVKMTHERLAL 1024
Db 335 NKHLELQSLNSKLGDSGIELQALSQPGTTLRLVLGLDCEVTNSGCSLASL 388

RESULT 6
S27880
Nasopressin receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 05-Nov-1999
C:Accession: S27880
R:Herrera, V.L.; Ruiz-Opazo, N.
Submitted to the EMBL Data Library, February 1992
A:Description: Characterization of a cDNA encoding an AII and AVP receptor isoform.
A:Reference number: S27880
A:Accession: S27880
A:Molecule type: mRNA
A:Residues: 1-483 <HEP>
A:Cross-references: EMBL:M85183; NID:G202805; PIDN:AAA03623.1; PID:G202806
C:Genetics:
A:Gene: AVP

Query Match 9.0%; Score 494.5; DB 2; Length 483;
Best Local Similarity 28.8%; Pred. No. 2.9e-30;
Matches 147; Conservative 74; Mismatches 205; Indels 85; Gaps 10;

Qy 391 LEGGGLLQRTGRTTAVTMLYLLSLMQPK-PGAPRLQPPNQRCGLCSLAADGLWNQKILF 449
Db 1 MELGRDLRSTKTTSTVLLFTITSLKLSAGTNGPRVQ--GELRMCLRLAREGLIKHQAP 58
Qy 450 EQDRL--KHGLDGDV-SAFILNMNI FOKDNCERYYSFIHLSFQEFPAAMYIILDEGEG 506
Db 59 SEKDRLERLKLQGSQVQTMFLSKKELPGVLETVVTVYQFIDQSQEFILAALSYLLD-AEG 117
Qy 507 GAGPDQDVTLLTEYAFERSFLATSRFLPGLNNEETRSHLEKSLCWKSPHIMDLQ 566
Db 118 AFGNSAGSVQMLNSDAGLRHLALTTRFLGLLSTERIDRIGNHFGCVVPRVYKQDTLR 177
Qy 567 WIOSKAQSDGSTL-----QOGLSREFFSCLYIIEQEEFIQ 601
Db 178 WVGOSQPKVATVGAKEKDELKDEABEEREEELNFGLELLYCLYETQEDDFVQ 237
Qy 602 ALSHFQVIVVSNI-ASKMEHVVSSFCRLKCRSAQVHLHYGATYSADGEDRAR----- 652
Db 238 ALSSUPENVLERVLRTRMDLEVLVYCVQCCPDGQALRLVSCGLVAAKEKKKKKSFNRL 297
Qy 653 ---CSAGATLLVQLRPVTLVDAYSEHAAALCTNPNLIELSLRYNALSGRGVKLLCQ 709
Db 298 KGSQSTGKOPPASLLRP-----LCE 317
Qy 710 GLRHPNCKLONLRKCRISSEACEDLSAALIANKNLTRMDLSGNGVGFPGMMLICEGLR 769
Db 318 AMITQOGLSILTLSHCKLPDVCARDLSEALKVAPSLRELGLLQNLRLTEAGLRLLSOGLA 377
Qy 770 HPQCRLOMIQLRKCQLESACQEMASVLGNTNPHLVELDLTGNALEDGLRLLCQGLRHPV 829
Db 378 WPKCKVQTLRIQMPGLQE-VIHYLVIVLQQSPVLTLLDLSGCLPFTVVEPLCSALKHPK 436
Qy 830 CRLRTWLKICRLTAACDELASTLSVNQSL 860
Db 437 CGLKTLSLTSLTENVLENPLRELQAVKTLKPDLL 467

RESULT 7
A48843
MHC class II transactivator - human
N:Alternate names: CIITA
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
R:Steinme, V.; Otten, L.A.; Zufferey, M.; Mach, B.
C:Accession: A48843
Cell 75, 135-146, 1993
A:Title: Complementatation cloning of an MHC class II transactivator mutated in hereditary
A:Reference number: A48843; MUID:94006536; PMID:8402893
A:Accession: A48843
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-1130 <STE>
A:Cross-references: UNIPROT:P33076; GB:X74301; NID:G414112; PIDN:CAA52354.1; PID:G41411

Query Match 6.9%; Score 377.5; DB 2; Length 1130;
Best Local Similarity 23.6%; Pred. No. 1.2e-20;
Matches 220; Conservative 122; Mismatches 355; Indels 235; Gaps 40;

Qy 92 EDLVRDPQETY-----RDYVRKRLMEDNARILGECVNLSHRYTRLLLVKEHSPNQV 145
Db 339 EQFYRLQDTYCAEPAGPDGILVEVLQV---ARL-----ERSSSKSLERELATPDWA 388
Qy 146 QOQLLDTGRGHARTVGHQASPIKTIETLPEDEERPEPRTVVMQAAAGIGKSLAHKVM 205
Db 389 ERQLAQGGAEV-----LLAAKEHRRPRETRVIAVLGKAGQGSYWAGAVSR 435
Qy 206 DWADGKLFQGRFDYLFYINCREMOSATECSMODLIFSCWPEPSAPLOE----LIRVPER 261
Db 436 ANACGRLPQ--YDFVFSVPCHLNRPDAYGLQDLFLSLGQPLVAADEVFSLHKRPDR 493
Qy 262 LFLFDGDELKPS---HDPQGPWCLWEKKRPEL--LLNSLRKLLKLLPELSLLITTR 316
Db 494 VLLIILDAFEEEAQOGFLHSTCGP-----APAPCSLRLGLLAGLFQKLLRGCTILLTAR 548
Qy 317 P-----TALEKLRHLLHPRHVEILGFSEAEERKEYFYKFHNA-----EQAGQVNTVRONE 368
Db 549 PGRVLVQSLSKADALF-----ELSGFSMEQAQAVVMRYFSSGTMTEHQDRALTLLDRP 602
Qy 369 PLFTMCVPLVWVCTCLQOOLBEGGLLRQTSRTTAVTMLYLLSLMQPKGAPRLQPP 428
Db 603 LLLSHSHSTPLCRVQCQLSEALLELGEDAKLPS-TLTGLYVGLL-----GRAALDSP 653
Qy 429 PNQRG-LCSLAAD-GLWNOKILFEEQ-----DLR-----KHGLDGEDVSAFLNMNIFQKIN 478
Db 654 PGALAEALAKLAWELGRRHQSTLQEDQFPADRVRTWMAKGLVQHPPRA----- 701
Qy 479 CERYYSFIHLSFQEFPAAMYIILDEGEGAGPDQDVTLLTEYAFER---SFLALTSEF 535
Db 702 AESELAPPSFLQCFLGALWAL---SGEIKDELPOYLALTPRKRPYDNWLEGVPRF 757
Qy 536 LFGILNEETRSHLEKSLCWKSPHIMDLQIQTQS---KAQSDGSTLQOGLSREFFSCLYE 592
Db 758 LAGLIFQPPARCIGALL--GPSAAASVDRKQKVLARYLRLQPGTILRAQLLELLHCAHE 815
Qy 593 IQEEEFQQALSHFQIVVSNIAKMEHVVSSFCRLKCRSAQVHLHYGATYSADGEDRAR 652
Db 816 AEAAGIQH-----VVQELPGRLL-----SFLGTRLTTPDA-HVLGKALEAAGQD--- 858
Qy 653 CSAGATLLVQLRPVTLVDAYSEHAAALCTN--PNLIELSL---YRNALSGRGVKLL 707
Db 859 -----FSLDLR-----STGICPSGLSGVLGSCVTRFRALSLD--TVAL 895
Qy 708 COGLR-HPNCKL-----QNL-RLKRCRISSEACEDLSAALIANK 744
Db 896 WESLRQIGETKLLQAABEKEFTIEFPKAKSLKDVEDLGLVQVQTRTSSSEDTAGELPAVR 955
Qy 745 NLTRMDLSGNGVGFPGMMLICEGLRHQPQCRLOMIQLRKCQLESACQEMASVLTGTHPLV 804
Db 956 DLKKLEFALGFVSGPQAF-----PKLV 977


```

RESULT 9
T31668
hypothetical protein COSI.5 - sea squirt (Ciona intestinalis)
C:Species: Ciona intestinalis
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31668
R:Bird, A.P.; Clark, V.; Jones, S.J.; Leitgeb, S.; Lennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z21050
A:Accession: T31668
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1075 <BIR>
A:Cross-references: UNIPROT:Q94430; EMBL:Z80904; PIDN:CAB02589.1
C:Geneticb:
A:Introns: 48/1, 557/3, 611/3, 685/2, 721/2, 739/1, 779/3, 820/3, 853/2, 914/2

Query Match 4.0%; Score 216.5; DB 2; Length 1075;
Best Local Similarity 20.8%; Pred. No. 3.5e-08;
Matches 153; Conservative 112; Mismatches 295; Indels 177; Gaps 29

Qy 102 YRDYVRKFRLEMDRNBALGECVNLSHRYTRLLLVKHSNPNMQOQLLDTGRGHARTVG 161
: || : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 HRD---EKHFVLQARAHATTGNGVFGSQVDVTIHQHPAPVNVQPL----- 47

Qy 162 HQASPIKIETLFE-----PDERPEPR-----TVMQGAAGICKSMLAHKVM 204
: || : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 48 -----VTIDQLFDKALENAQRYTEDESAEYAKWYIERHANTVWVGPPGVGKTTLL-KKM 101

Qy 205 LDWADGKLFQGRFDYLFVINCENMQSATECSMDLI-----PSCWPEPSAPLQELIRV 258
: : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 102 VKQILKHEFLPDTYIFFINVKDIDFNKEMTLLEFLTNSRVKNVNTBEESKALITFLN 161

Qy 259 PERLLFIIDGDELKPSPHDQGPWCCLCWECKRPPELLNSLRKCLAPELSLIITRPT 318
: || : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 162 NPNVAIFDGDJDEASTNEFARIPHIKLDGKSKPVDIMKN-LFNLTLLPKAKIVTSTLH 220

Qy 319 ALEKHLRLLEHPHVEILGFSAEERKEYFYKFNHSAQAGQVFNTYR--DNEP-LFTWC 374

```

Db 221 QMKLHPDYRPTSI FEVLGLLEAKNNL-----GTQLCGEYPAIKTLDQOPNLAHLC 274
QY 375 FVPLVWVWCTCLQQLEGGLLRQTSRTTAVYMLYLLSLMQ-----PKFGAPR 424
Db 275 YLPFINFLVFCLLSN-EGSDI-----KWTQVLIFSWTRFVELSHLKGSEVLDKVGAE 328
QY 425 LQPPPNQRGICSLAADGLMNQKILFBEQDLRKHGLDGE-----DVSAPLNNITFQ 474
Db 329 VK-----LARLAYKGLQKLVFEKTFDDVKLADVMVTFNPFHTYVDKSGIRMKILE 381
QY 475 KQNCERYYSFIHLSFOEFAAMYIILDEGEGAGPDQDVTRLLTYAERSSEFAL---- 531
Db 382 GN---KRSY-FTHLIWQEFYAAYLML-----FVSREFEOLKEPI 417
QY 532 -----TSRELFGLLNEETRSHLEKSLCWKYSVP-----HIKMDLLQWIOSKAQSDG 576
Db 418 FKDAQKRVVGFMPGICNPAYQOLK-----LVFPATMIKDYEEKELM--VPMWESLW 470
QY 577 STLOQGSLEFFSCLYIEQEETIOQALSHFOV---IVVSNIAKMEHMYSSFCLEKCRSA 633
Db 471 SARGEDLIRRFGLWHEYNDESSKKFEDYLPVGLKMDAPKHLSEVKDLV--YALKSFTRP 528
QY 634 QVLHL--YGATYSADGEDRARCAGATLIV-----QLRP 666
Db 529 HKLRLDSYETTTEVLETLRLRGVHGTTTTITRFVNNIEMKDSIMELLLLHLDAEELRF 588
QY 667 ERTVLLDAYSEHLAAALCTPNPILIELSLYRNAL--GSRGVKLL--COGLRHPNCKLQNLRL 723
Db 589 IRVTNLSFYMERLSNAINQSRNKIQLVLIKHLHDDYDVYKLAGCLG-----NISLYM 642
QY 724 KCRISISSACEDLSAAL 740
Db 643 WGTDISSDQCSVLKQAI 659

RESULT 10
A55478
Neuronal apoptosis inhibitory protein - human
N/Alternate names: NAIP
C/Species: Homo sapiens (man)
C/Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 02-Feb-2001
C/Accession: A55478
R/ROY, N.; Mahadevan, M.S.; McLean, M.; Shuttler, G.; Varaghi, Z.; Parahani, R.; Baird, S.
d, T.O.; de Jong, P.J.; Surh, L.; Ikeda, J.E.; Korneluk, R.G.; Mackenzie, A.
Cell 80, 167-178, 1995
A/Title: The gene for neuronal apoptosis inhibitory protein is partially deleted in indi
A/Reference number: A55478; MUID:95112344; PMID:7813013
A/Accession: A55478
A/Molecule type: mRNA
A/Residues: 1-1232 <ROY>
A/Cross-references: GB:U19251
C/Genetics:
A/Gene: GDB:SMA@; SMA
A/Cross-references: GDB:120378; OMIM:600354; OMIM:253300
A/Map position: 5q12.2-5q13
C/Keywords: apoptosis; ATP; glycoprotein; nucleotide binding; P-loop; transmembrane prot
F:94-110/Domain: transmembrane #status predicted <TM1>
F:470-477/Region: nucleotide-binding motif A (P-loop)
F:479-496/Domain: transmembrane #status predicted <TM2>
F:476/Binding site: ATP (lys) #status predicted
F:618,632,823,923,1035/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 3.4%; Score 183.5; DB 2; Length 1232;
Best Local Similarity 21.3%; Pred. No. 1.6e-05;
Matches 148; Conservative 94; Mismatches 248; Indels 205; Gaps 35;

QY 188 MOGAGIGKSLAHKLVMDWADGKL-FQGRFDLYFINCR-----EMN 229
Db 468 VGEAGSGKTVLLKIAFLWASGCCPLLNRFLVFLVLSLRPRDEGLASITCDQLLEKE 527
QY 230 QSATESCMQDLIFSCWPEPSAPLQELIRVPERFLIIDGFDLKPSPHDPQPCWCLWEE 289

Db 528 GSVTEMCWRNII-----QQI---KNQVLFLLDDYKEI---CSIQP----- 561
QY 290 KRPTELLNSLRKLLPELSLLITTRPTALEKLRHLEHPRHVILGSEAEKKEYFYK 349
Db 562 -----VIGKLIQKHLSTRCLLIHAVTRNARDIRRYLE--TILEIQAFP-----FYN 606
QY 350 -----YFNASQAGOVFNVRDNEPLFTMCVPVPLVCVVVCT-CLQQOLEGGGLLRQT 400
Db 607 TVCILRKLFSSWTRLRKFWYFGKNSLOKIQKTPLFVAACAHWFQYFPD-----P 659
QY 401 SRTTAVTMYLLSL-MOPKGPAPLQPPPNQRGICSLAADGLMNQKILFBEQDLRKHGL 459
Db 660 SFDDVAVPKSYMERLSLRNKATAILKATVSSCG--ELALKGFSCCFEFNDDDLAAGV 717
QY 460 DGEDVSAPLNNI FQKQDINCERYYSFIHLSFOEFAAMYI--LDEGEGAGPDQDV--- 514
Db 718 DEDEDLTVCLMSKFTAQ-RLRPYRFUSPAPQEFGLAGMRLIELLDSR--QEHQDLGLY 773
QY 515 -----TRLLTYAF-----SERSFLATLSRFLGGLNEETRSHL----- 548
Db 774 HLKQINSMPMTVSAXNNFLNVSSLPSTKAGPKIVSHLLHLVDNKNESLENISENDDYLKH 833
QY 549 --EKSL-----CWKSPPHIKMDL-----LQWIOSKAQ 573
Db 834 QPEISLQMLLRGLMQICFOAYFMSVSEHLLVLAALKTAQSNVTVAACSPFVLQFLQ----- 889
QY 574 SDGSTLOQGSLEFFSCLYIEQEETIOQALS-HFOVIVVSNIAKMEH-VSSFCLEKRCR 631
Db 890 ---GRTLTGAL---NLQYFFDPESLSLRSIHFS--IRGNKTSAPRAHFSVLETCFDSQ 942
QY 632 SAQVHLXYCATYS-----ADGED-----RARCAGATLIVLQRPVTL--- 671
Db 943 VPTIDODYASAPEPMNEWERNLAEKEDNVKSYMDMQRASPDLSYGYWKLSPKQYKIPCL 1002
QY 672 -----LDAYSEHLAAALCT---NPNLIELSLYRNALSGRGVKKLCOGLRHPNCKLQNLRL 722
Db 1003 EVDVNDIDVVGQDMLEILMTVFSASQRIELHNLH---SRG---FIESIR-PALELSKAS 1054
QY 723 LKCRISISSACEDLSAA-----LIANKNLTRMDLSG 753
Db 1055 VTKCSISKL---ELSAABEQELLLTLPSLESLEVSQ 1086

RESULT 11
B97746
hypothetical protein RC0370 [imported] - Rickettsia conorii (strain Malish 7)
C/Species: Rickettsia conorii
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C/Accession: B97746
R/Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
Science 293, 2093-2098, 2001
A/Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A/Reference number: A97700; MUID:21442074; PMID:11557893
A/Accession: B97746
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-312 <KUR>
A/Cross-references: UNIPROT:Q921Q0; GB:AE06914; PIDN:AA02908.1; PID:g15619435; GSPDB:
C/Genetics:
A/Gene: RC0370

Query Match 3.3%; Score 180; DB 2; Length 312;
Best Local Similarity 19.0%; Pred. No. 4e-06;
Matches 64; Conservative 74; Mismatches 117; Indels 82; Gaps 12;

QY 97 DPQETVDYVRRKFRMLMDRNLGBCVNLSHRY-----TRLLVKEHNPQM 144
Db 34 DTEAKTREWFDKYNQVEEKIDNIAKSLNEENKFAILLRNLTWNKSTSEIKNIVEKINRMQ 93
QY 145 VQOQLDGTGRHARTVGHQASPIKIEITLFEDEPERPEPPTVYMOGAGIGKSLAHKVM 204
Db 94 V-----DIGK-----VLLGSAGIGKTTLMHYLS 117

| | | | |
|-----|--|----------|-----|
| 930 | WGTLWLLAEGLOHPACRKLQKLMWDSC-GLTAKACENLYFTLGINTQLTDL--- | YLTNNALG | 985 |
| | : | : | : |
| 426 | VGIStIAQGCIH-----LETINISYCODITDKSLVSL-SKCSLLGTFSRGCPNIIISOGLA | 480 | |
| | : | : | : |
| | : | : | : |
| | : | : | : |
| 986 | DTGVRLLLCKRLSHPGCK-----LRVLMLFGMDLNMKTHSKRLAALRVTKPYL-DIGC | 1035 | |
| | : | : | : |
| | : | : | : |
| 481 | AIAVR--CKELAKVDILKKCFPSINDAGLLALAHLFONLKQINVSTAVTEVGLIUSLANIGC | 538 | |
| | : | : | : |
| | : | : | : |
| | : | : | : |

Search completed: February 8, 2005, 14:59:56
Job time : 57 secs

RESULT 15

T48193
Hypothetical protein F7A7.240 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T48193
R:R.Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 2000
A:Reference number: 224487
A:Accession: T48193
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-618 <BV>
A:Cross-references: UNIPROT:Q9M004; EMBL:AL161946
A:Experimental source: cultivar Columbia; BAC clone F7A7
C:Genetics:
A:Map position: 5
A:Introns: 213/3; 255/3; 259/3; 294/3; 350/3; 376/1; 419/2; 517/3; 601/3
A:Note: F7A7.240

| | | | | |
|-----------------------|------------------|--|----------|-------------|
| Query Match | 3.2%; | Score 172.5; | DB 2; | Length 618; |
| Best Local Similarity | 25.4%; | Pred. No. 4.1e-05; | | |
| Matches 122; | Conservative 61; | Mismatches 137; | Gaps 28; | |
| QY | 629 | RCRSQVHLHGATYSADGEDRCARSAGATHLVLQVLRPERTVLLDAYSEHIAAALCTNP | 688 | |
| DB | 123 | KCVNLVEIDLSNATEMRDAD--AAVAAEARS--LERLUGRCRKMLTDMGIGCTAVCKKLN | 179 | |
| QY | 689 | LLELSLYRNALGSRGVKLL--COGLR-------HPNCKLQNLR---LKRC-- | 726 | |
| DB | 180 | TVSLK--WCVGVGDVLGVGLAVKCKDIETLDLSVLPITGKCLHDILKIQHLELLLEGCGF | 238 | |
| QY | 727 | -----RISSACEDLSAALTANKNLTRMDLSGNGVGFPGWMLLC | 765 | |
| DB | 239 | VDDDSLKSLRHDCSKLWKYQKLDASSQNLT-----HRGLTSL--LSGAGY----- | 283 | |
| QY | 766 | EGLRHPQCRQLQWILQRKQLESGACQEWASVLGTNPHLVELDITGNALEDIGLR--- | 822 | |
| DB | 284 | -----LQRLDLSHC--SSVLSLDFASLSKVKVSALOSIRJDDGCVTPDGLKKAIGTLC | 332 | |
| QY | 823 | QGLRH-----PVCRLATLML--KICRLTAAACDELASTLSVNQSLRELDLSNKLGD | 873 | |
| DB | 333 | NSLKEVSLSKCVSVTDEEAFMLIGQKRL-----LEELDITDNEIDD-- | 374 | |
| QY | 874 | GVLLLCGLRH--PTC--KLQTLRLGICRIGSAACEGLSVLQVLANHNRLDELDFNDIG-- | 929 | |
| DB | 929 | -----PCLKSTSSGTSLSSTKATCT--NTTDDKHSVIGWGCNSLRLELDL--YRSVGITD | 425 | |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 14:29:33 ; Search time 185 Seconds
(without alignments)
2864.875 Million cell updates/sec

Title: US-10-781-294-24

Perfect score: 5472

Sequence: 1 MLRTAGRDGLCRSLSTYLEEL.....NTHSLAALRVTKPYLDIGC 1035

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

.Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 5433 | 99.3 | 1061 | 1 | NA12 HUMAN |
| 2 | 2722 | 49.7 | 719 | 2 | Q6UQE6 |
| 3 | 2396 | 43.8 | 1034 | 1 | CS1_HUMAN |
| 4 | 2359.5 | 43.1 | 1033 | 1 | CS1_MOUSE |
| 5 | 2253.5 | 41.2 | 977 | 2 | Q6TG35 |
| 6 | 1799 | 32.9 | 1093 | 1 | NA14_HUMAN |
| 7 | 1519.5 | 27.8 | 991 | 2 | Q7RTR0 |
| 8 | 1505.5 | 27.5 | 986 | 2 | Q86W27 |
| 9 | 1493.5 | 27.3 | 506 | 2 | Q68U61 |
| 10 | 1475.5 | 27.0 | 506 | 2 | Q68U59 |
| 11 | 1475.5 | 27.0 | 506 | 2 | Q68U60 |
| 12 | 1474 | 26.9 | 500 | 2 | Q68U56 |
| 13 | 1471 | 26.9 | 500 | 2 | Q68U53 |
| 14 | 1470.5 | 26.9 | 506 | 2 | Q68U57 |
| 15 | 1468.5 | 26.8 | 506 | 2 | Q68U58 |
| 16 | 1461 | 26.7 | 500 | 2 | Q68U51 |
| 17 | 1458 | 26.6 | 500 | 2 | Q68U50 |
| 18 | 1457.5 | 26.6 | 994 | 1 | NA14_HUMAN |
| 19 | 1450.5 | 26.5 | 994 | 2 | Q68W87 |
| 20 | 1449.5 | 26.5 | 499 | 2 | Q68U55 |
| 21 | 1442 | 26.4 | 500 | 2 | Q68U54 |
| 22 | 1438 | 26.3 | 500 | 2 | Q68U52 |
| 23 | 1402 | 25.6 | 475 | 2 | Q68U48 |
| 24 | 1401 | 25.6 | 1043 | 1 | NA13_HUMAN |
| 25 | 1396.5 | 25.5 | 993 | 2 | Q6B966 |
| 26 | 1390.5 | 25.4 | 986 | 2 | Q66X14 |
| 27 | 1387 | 25.3 | 1098 | 2 | Q64719 |
| 28 | 1377 | 25.2 | 1375 | 2 | Q6CUB5 |
| 29 | 1375 | 25.1 | 1473 | 1 | NA11_HUMAN |
| 30 | 1320 | 24.1 | 1200 | 1 | NA15_HUMAN |
| 31 | 1314.5 | 24.0 | 287 | 2 | Q9BY26 |

| | | | | | |
|----|--------|------|------|---|------------|
| 32 | 1309 | 23.9 | 1062 | 1 | NA12_HUMAN |
| 33 | 1304 | 23.8 | 982 | 2 | Q66X27 |
| 34 | 1299.5 | 23.7 | 437 | 2 | Q68U49 |
| 35 | 1270.5 | 23.2 | 413 | 2 | Q6JEK7 |
| 36 | 1268 | 23.2 | 978 | 2 | Q66X19 |
| 37 | 1259 | 23.0 | 982 | 2 | Q8BU40 |
| 38 | 1251.5 | 22.9 | 413 | 2 | Q6JEK9 |
| 39 | 1249.5 | 22.8 | 1048 | 2 | Q7RTR4 |
| 40 | 1222 | 22.3 | 1009 | 2 | Q7RTR1 |
| 41 | 1216.5 | 22.2 | 1029 | 2 | Q86W28 |
| 42 | 1183 | 21.6 | 949 | 2 | Q66X03 |
| 43 | 1168.5 | 21.4 | 1111 | 2 | Q7TPU9 |
| 44 | 1166.5 | 21.3 | 1095 | 2 | Q6VSG3 |
| 45 | 1165.5 | 21.3 | 1111 | 2 | Q6XZF2 |

ALIGNMENTS

RESULT 1
NA12_HUMAN
ID NA12_HUMAN STANDARD; PRT; 1061 AA.
AC PS9046; OSNEU4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE NACHT-, LRR- and PYD-containing protein 12 (PYRIN-containing APAF1-like protein 7) (Monarch-1).
GN Name=NALP12; Synonyms=PYPAF7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22162427; PubMed=12019269; DOI=10.1074/jbc.M203915200;
RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S., Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates activation of NF-kappa B and caspase-1-dependent cytokine processing.";
RT J. Biol. Chem. 277:29874-29880(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22451042; PubMed=12563287; DOI=10.1038/nrml019;
RA Tschopp J., Martinon F., Burns K.;
RT "NALPs: a novel protein family involved in inflammation.";
RT Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4).
RX TISSUE=Lymphoma;
RA Williams K.L., Linhoff M.W., Harton J.A., Ting J.P.Y.;
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX TISSUE=Leukocyte;
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshul S.F., Zeng H., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Roh S.S., McQuellano N.A., Peters G.J., Abramson R.D., Mullihy S.J., Bhat N.K., McKernan K.J., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,

RA Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences";
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
 CC activation of NF-kappa-B via IKK.
 CC -!- SUBUNIT: Binds to ASC with its DAPIN domain.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=4;
 CC Names=1; Synonyms=1;
 CC IsoId=P59046-1; Sequence=Displayed;
 CC Name=2; Synonyms=11;
 CC IsoId=P59046-2; Sequence=VSP_005524;
 CC Name=3; Synonyms=111;
 CC IsoId=P59046-3; Sequence=VSP_005523;
 CC Name=4; Synonyms=IV;
 CC IsoId=P59046-4; Sequence=VSP_009879;
 CC -!- TISSUE SPECIFICITY: Detected only in peripheral blood leukocytes,
 CC predominantly in eosinophils and granulocytes, and at lower levels
 CC in monocytes.
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 leucine-rich (LRR) repeats.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AY095146; AAM18227.1; -
 DR EMBL; AY154467; AAO18163.1; -
 DR EMBL; AY116204; AAM75142.1; -
 DR EMBL; AY116205; AAM75143.1; -
 DR EMBL; AY116206; AAM75144.1; -
 DR EMBL; AY116207; AAM75145.1; -
 DR EMBL; BC028069; AAM75145.1; -
 DR HSSP; P13489; 1A4Y.
 DR Genew; HGNC:22938; NALP12.
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0008656; F:caspase activator activity; NAS.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR GO; GO:0006919; P:caspase activation; NAS.
 DR GO; GO:0050729; P:positive regulation of inflammatory response; NAS.
 DR GO; GO:0050718; P:positive regulation of interleukin-1 beta s...; NAS.
 DR GO; GO:0045381; P:regulation of interleukin-18 biosynthesis; NAS.
 DR GO; GO:0008588; P:release of cytoplasmic sequestered NF-kappaB; IDA.
 DR GO; GO:0007165; P:signal transduction; NAS.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN.
 DR Pfam; PF00560; LRR; 7.
 DR Pfam; PF05729; NACHT; 1.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PRINTS; PR00019; LEURICHRPT.
 DR PROSITE; PS50834; DAPIN; 1.
 DR PROSITE; PS50837; NACHT; 1.
 KW Alternative splicing; ATP-binding; Leucine-rich repeat; Repeat.
 FT DOMAIN 1 95
 FT DAPIN.
 FT DOMAIN 211 528
 FT REPEAT 712 735
 FT REPEAT 741 764
 FT REPEAT 769 791
 FT REPEAT 798 821
 FT REPEAT 826 849
 FT REPEAT 883 906
 FT REPEAT 940 968
 FT REPEAT 997 1020
 FT NP_BIND 217 224
 FT ATP (Potential).

| | | | | |
|-----------------------|--------------|--|--------------|-----------------------------|
| FT | VARSPLIC | 862 | 973 | Missing (in isoform 3). |
| FT | VARSPLIC | 976 | 1031 | /FTid=VSP_005523. |
| FT | VARSPLIC | 862 | 1031 | Missing (in isoform 2). |
| FT | VARSPLIC | 862 | 1031 | /FTid=VSP_005524. |
| FT | VARSPLIC | 862 | 1031 | Missing (in isoform 4). |
| FT | CONFLICT | 691 | 691 | /FTid=VSP_009879. |
| FT | SEQUENCE | 1061 | AA; 120172 | MM; 8C10AFB4907C131B CRC64; |
| Query Match | | 99.3% | Score 5433; | DB 1; Length 1061; |
| Best Local Similarity | | 97.4% | Pred. No. 0; | |
| Matches 1034; | Conservative | 0; | Mismatches | 0; Indels 28; Gaps 2; |
| QY | 1 | MURTAGRDLGRLSTYLLLEAEVAVLKKFKLYLGTATLGEKGIIPWGSMEKAGPLEMAQLL | 60 | |
| DB | 1 | MURTAGRDLGRLSTYLLLEAEVAVLKKFKLYLGTATLGEKGIIPWGSMEKAGPLEMAQLL | 60 | |
| QY | 61 | ITHFGPEEAWRLALSTPERINRKLWEGQREDLVR----- | 96 | |
| DB | 61 | ITHFGPEEAWRLALSTPERINRKLWEGQREDLVR----- | 120 | |
| QY | 97 | ----DPQETRYDYVRKFKLMEDRNALGECVNLSHRYTRLLLVKHSNPMQVQQLDGTG | 153 | |
| DB | 121 | PRKDPQETRYDYVRKFKLMEDRNALGECVNLSHRYTRLLLVKHSNPMQVQQLDGTG | 180 | |
| QY | 154 | RGHARTVGHQASPIKIETLFFPDEERPEPRTVVMQGAAGIGKMLAHKVMLDWADGKLF | 213 | |
| DB | 181 | RGHARTVGHQASPIKIETLFFPDEERPEPRTVVMQGAAGIGKMLAHKVMLDWADGKLF | 240 | |
| QY | 214 | QGRFDYLFVINCENQATSCMODLIFSCWPEPSAPLOELIRVPERLLFTIDGFDLX | 273 | |
| DB | 241 | QGRFDYLFVINCENQATSCMODLIFSCWPEPSAPLOELIRVPERLLFTIDGFDLX | 300 | |
| QY | 274 | PSFHDPOGQWCLCWEKRPTELLNSLRKLLPELSLITTRPTALEKHLRLHPRHV | 333 | |
| DB | 301 | PSFHDPOGQWCLCWEKRPTELLNSLRKLLPELSLITTRPTALEKHLRLHPRHV | 360 | |
| QY | 334 | EILGFSEAEKRYFYKYFHNAGQGVFNVDNPEPLFTMCVFPLVWVWCTCLOOQLG | 393 | |
| DB | 361 | EILGFSEAEKRYFYKYFHNAGQGVFNVDNPEPLFTMCVFPLVWVWCTCLOOQLG | 420 | |
| QY | 394 | GGILROTSRTTAVVMYLLSLMOPGAPRQPPNORGLCSLAADGLWNQKILFEED | 453 | |
| DB | 421 | GGILROTSRTTAVVMYLLSLMOPGAPRQPPNORGLCSLAADGLWNQKILFEED | 480 | |
| QY | 454 | LRKHGLDGEDVSAFLNMNIFQKDINCERYYSFIHSFQFFFAAMYIILDEGEGGAPDOD | 513 | |
| DB | 481 | LRKHGLDGEDVSAFLNMNIFQKDINCERYYSFIHSFQFFFAAMYIILDEGEGGAPDOD | 540 | |
| QY | 514 | VTLLTEYAFSERSFLATSRFLFGLNEETSHLEKSLCWKVSPIHKMDLLQWIOSKAQ | 573 | |
| DB | 541 | VTLLTEYAFSERSFLATSRFLFGLNEETSHLEKSLCWKVSPIHKMDLLQWIOSKAQ | 600 | |
| QY | 574 | SDGSTLQOQSLEFFSCLYEIEEFTQQALSHFQVIVVSNIAKMEHMYSSFCLEKCRSA | 633 | |
| DB | 601 | SDGSTLQOQSLEFFSCLYEIEEFTQQALSHFQVIVVSNIAKMEHMYSSFCLEKCRSA | 660 | |
| QY | 634 | QVLHLYGATYSADGEDRARCAGAHILLVOLPERVTLDAYSEHLAAALCTNPNIELS | 693 | |
| DB | 661 | QVLHLYGATYSADGEDRARCAGAHILLVOLPERVTLDAYSEHLAAALCTNPNIELS | 719 | |
| QY | 694 | LYRNALSGRGVKKLCCGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSG | 753 | |
| DB | 720 | LYRNALSGRGVKKLCCGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSG | 779 | |
| QY | 754 | NGVGFPGMMLLCEGLRHPQCRLOMIQLRKQLESAGCQEMASVLTGNPHLVELDLTGAL | 813 | |
| DB | 780 | NGVGFPGMMLLCEGLRHPQCRLOMIQLRKQLESAGCQEMASVLTGNPHLVELDLTGAL | 839 | |
| QY | 814 | EDLGLRLCCGLRHPVCLRTLMKICRLTAAACDELASTLSVNQSLRELDLSINELGDL | 873 | |
| DB | 840 | EDLGLRLCCGLRHPVCLRTLMKICRLTAAACDELASTLSVNQSLRELDLSINELGDL | 899 | |

QY 874 GVLLCEGLRHPTCKLQTLRLGICRLGSAACEGLSVVLQAHNNRELDLSPNDLGDWGLW 933
 Db 900 GVLLCEGLRHPTCKLQTLRLGICRLGSAACEGLSVVLQAHNNRELDLSPNDLGDWGLW 959
 QY 934 LLAEGLOHPACRLQKLDSCGLTAKACENLYFTLGINQTLTDLTLTNALGDTGVRLIC 993
 Db 960 LLAEGLOHPACRLQKLDSCGLTAKACENLYFTLGINQTLTDLTLTNALGDTGVRLIC 1019
 QY 994 KRLSHPGCKRLVLMFGMDLNKMTSHRLAALRVTKPYLDIGC 1035
 Db 1020 KRLSHPGCKRLVLMFGMDLNKMTSHRLAALRVTKPYLDIGC 1061

RESULT 2
 Q6UQE6 PRELIMINARY; PRT; 719 AA.
 AC Q6UQE6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE NALP12 (Fragment)
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY364010; AAQ03590; LRR_Rninh sub.
 DR InterPro; IPR003590; LRR_Rninh sub.
 DR InterPro; IPR007111; NACHT NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN.
 DR Pfam; PF05729; NACHT; 1.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR SMART; SM00368; LRR_RI; 1.
 DR PROSITE; PS00824; DAPIN; 1.
 DR PROSITE; PS00837; NACHT; 1.
 FT NON TER 719
 SQ SEQUENCE 719 AA; 82623 MW; CA2E777CE7CC4A1F CRC64;

Query Match 49.7%; Score 2722; DB 2; Length 719;
 Best Local Similarity 71.1%; Pred. No. 3.8e-185;
 Matches 517; Conservative 81; Mismatches 115; Indels 14; Gaps 3;

QY 1 MLRTAGDGLCRSLTYLLEEAVALKKFKLYLGTATLGECKIPGWSMEKAGPLEMAQLL 60
 Db 1 MLPSTARDGLYRLTYLLEEAVALKKFKLYLGTATLGECKIPGWSMEKAGPLEMAQLL 60
 QY 61 ITHFGPEEAVRLALSTFERINRKLWERGQEDLVR-----DPQETRYDYVRKKFLME 114
 Db 61 VAHGTREAWLLALSTFQRHKKDLWERGQEDLVRGKGGDLQTTKYDYVRKKFLME 120
 QY 115 DRNARLGEVNLSHRYTRLLLVKEHSPMVOVQQLDTRGHARTVGHQASPIKIETLFE 174
 Db 121 DRNARLGEVNLNRYTRLLLVKEHSPNFIWQKRFVDVWEVSRRTRRQTSPIQMELFE 180
 QY 175 PDEERPPRTVVMQAGAGTKSLAHKVMLDWADGKLFQGRPDYLYINCREMNSQATE 234
 Db 181 PDEERPPRTVVMQAGAGTKSLAHKVMLDWADGKLFQGRPDYLYINCREMNSHTQ 240
 QY 235 CSMODLFFSCWPESAPLOELIRVPERLLFTIDGFDLKPFSFHPDQGPWCLCWEKRPTE 294
 Db 241 CSVDLISSCWPERGISLEDLMQAPDRLLFTIDGFDKLPFSFHPDQGPWCLCWEKRPTE 300
 QY 295 LLNSLRKXLLPESLLITRPTALBKLHRLBHPHRLVEILGFSPEARKEYFYKYPHNA 354
 Db 301 VLLGSLIRLLPQVSLITRPTALEKLLHGLBHPHRLVEILGFSPEARKEYFYRYPHNT 360
 QY 355 EQAQGVNRYRDNPEFTMCFVPLVCVWCTCLOQLEGGGLLRQTSRTTAVYMLYLS 414
 Db 361 QQASRVLSFLMDYEPFTMCFVPMVSVWVCTCLKQQLSGLLRQTSRTTAVYMYLLS 420

RESULT 3

CIS1_HUMAN

ID CIS1_HUMAN STANDARD; PRT; 1034 AA.
 AC Q36P20; O75434; Q8TCW0; Q8TEU9; Q8WXH9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Cold autoimmune syndrome 1 protein (Cryopyrin) (NACHT-, LRR- and
 DE PYD-containing protein 3) (PYRIN-containing APAF1-like protein 1)
 DE (Angiotensin/vasopressin receptor AII/AVP-like).
 GN Name=CIS1; Synonyms=NALP3, PYPAF1;
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), VARIANTS FCAS MET-198; VAL-439
 RP AND GLY-627, AND VARIANT MWS VAL-352.
 RX MEDLINE=21547523; PubMed=11687797; DOI=10.1038/ng756;
 RA Hoffman H.M., Mueller J.L., Broide D.H., Wenderer A.A., Kolodner R.D.,
 RT "Mutation of a new gene encoding a putative pyrin-like protein causes
 RT familial cold autoinflammatory syndrome and Muckle-Wells syndrome."
 RL Nat. Genet. 29:301-305(2001).
 RN (2)
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21909508; PubMed=11786556; DOI=10.1074/jbc.M112208200;
 RA Manji G.A., Wang L., Geddes B.J., Brown M., Merriam S., Al-Garawi A.,
 RA Mak S., Lora J.M., Briskin M., Jurman M., Cao J., Distefano P.S.,
 RA Bertin J.;
 RT "PYPAF1: a PYRIN-containing APAF1-like protein that assembles with ASC
 RT and activates NF-kB."
 RL J. Biol. Chem. 277:11570-11575(2002).
 RN (3)
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), VARIANT MWS MET-198, AND
 RP VARIANTS FCAS/MWS TRP-260 AND PRO-305.
 RX MEDLINE=22141234; PubMed=12355493; DOI=10.1002/art.10509;
 RA Aganna E., Martinon F., Hawkins P.N., Ross J.B., Swan D.C.,
 RA Booth D.R., Lachmann H.J., Gaudet R., Woo P., Feigherly C.,
 RA Cotter F.E., Thome M., Hitman G.A., Tschopp J., McDermott M.F.;
 RT "Association of mutations in the NALP3/CiAS1/PYPAF1 gene with a broad
 RT phenotype including recurrent fever, cold sensitivity, sensorineural
 RT deafness, and AA amyloidosis."
 RL Arthritis Rheum. 46:2445-2452(2002).
 RN (4)
 RP SEQUENCE OF 391-1034 FROM N.A. (ISOFORM 1).
 RX TISSUE=Umbilical cord blood;
 MEDLINE=20499367; PubMed=11042152; DOI=10.1101/gr.140200;

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,
RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,
RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.,
RT "Cloning and functional analysis of cDNAs with open reading frames for
RT 300 previously undefined genes expressed in CD34+ hematopoietic
RT stem/progenitor cells.";
RL Genome Res. 10:1546-1560(2000).
[5]
RN VARIANT FCAS MET-198, VARIANTS MWS ASN-303; MET-348; THR-439 AND
RP ARG-569, AND VARIANT FCAS MWS TRP-260.
RX MEDLINE=21987640; PubMed=11922256;
RA Dode C., Le Du N., Cusset L., Letourneur F., Berthelot J.-M.,
RA Vaudour G., Meyrier A., Watts R.A., Scott D.G.I., Nicholas S.,
RA Granel B., Frances C., Garcia F., Edery P., Boulanguet S.,
RA Domergues J.-P., Delpech M., Grateau G.;
RT "New mutations of CIAS1 that are responsible for Muckle-Wells syndrome
RT and familial cold urticaria: a novel mutation underlies both
RT syndromes.";
RL Am. J. Hum. Genet. 70:1498-1506(2002).
[6]
RN VARIANTS CINCA ASN-303; SER-309; ARG-358; ASN-436; SER-573 AND
RP THR-662, AND TISSUE SPECIFICITY.
RX MEDLINE=2062555; PubMed=12032915;
RA Feldmann J., Prieur A.-M., Quartier P., Berquin P., Certain S.,
RA Cortis E., Teillac-Hamel D., Fischer A., de Saint-Basile G.;
RT "Chronic infantile neurological cutaneous and articular syndrome is
RT caused by mutations in CIAS1, a gene highly expressed in
RT polymorphonuclear cells and chondrocytes.";
RL Am. J. Hum. Genet. 71:198-203(2002).
CC -!- FUNCTION: May function as a potential inducer of apoptosis.
CC Interacts selectively with apoptosis-associated specklike protein
CC containing a CARD domain (ASC). This complex may function as an
CC upstream activator of NF-kappa-B signaling.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=2;
CC IsoId=Q96P20-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q96P20-2; Sequence=VSP_005520, VSP_005521;
CC Name=3;
CC IsoId=Q96P20-3; Sequence=VSP_005519;
CC -!- TISSUE SPECIFICITY: Expressed in blood leukocytes. Strongly
CC expressed in polymorphonuclear cells, undetectable or expressed
CC at a lower magnitude in B and T lymphoblasts, respectively. High
CC level of expression detected in chondrocytes. Low or no expression
CC in the other tissues tested.
CC -!- DISEASE: Defects in CIAS1 are a cause of familial cold
CC autoinflammatory syndrome (FCAS) [MIM:120100]; commonly known as
CC familial cold urticaria. FCAS is rare autosomal dominant systemic
CC inflammatory disease characterized by episodes of rash,
CC arthralgia, fever and conjunctivitis after generalized exposure to
CC cold.
CC -!- DISEASE: Defects in CIAS1 are a cause of Muckle-Wells syndrome
CC (MWS) [MIM:191900]; a rare autosomal dominant fever syndrome with
CC episodic urticaria, arthralgia, amyloidosis and progressive
CC sensorineural deafness.
CC -!- DISEASE: Defects in CIAS1 are the cause of chronic infantile
CC neurologic cutaneous and articular syndrome (CINCA) [MIM:607115];
CC also known as 'neonatal onset multisystem inflammatory disease',
CC or NOMID, a rare congenital inflammatory disorder characterized by
CC a triad of neonatal onset of cutaneous symptoms, chronic
CC meningitis, and joint manifestations with recurrent fever and
CC inflammation.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- CAUTION: Ref.4 sequence differs from that shown due to frameshifts
CC in positions 893, 918 and 926.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF410477; AAL33908.1; -
DR EMBL; AF427617; AAL33911.1; -
DR EMBL; AY051117; AAL12497.1; -
DR EMBL; AY051112; AAL12497.1; JOINED.
DR EMBL; AY051113; AAL12497.1; JOINED.
DR EMBL; AY051114; AAL12497.1; JOINED.
DR EMBL; AY051115; AAL12497.1; JOINED.
DR EMBL; AY051116; AAL12497.1; JOINED.
DR EMBL; AY056059; AAL12497.1; JOINED.
DR EMBL; AY056060; AAL12497.1; JOINED.
DR EMBL; AY051117; AAL12498.1; -
DR EMBL; AY051112; AAL12498.1; JOINED.
DR EMBL; AY051113; AAL12498.1; JOINED.
DR EMBL; AY051114; AAL12498.1; JOINED.
DR EMBL; AY051115; AAL12498.1; JOINED.
DR EMBL; AY051116; AAL12498.1; JOINED.
DR EMBL; AF420469; AAL65136.1; -
DR EMBL; AF468522; AAL78632.1; -
DR EMBL; AY092033; AAM14669.1; ALT_INIT.
DR EMBL; AF418985; AAL14640.2; ALT_INIT.
DR EMBL; AF054176; AAC39910.1; ALT_FRAME.
DR HSP; P10775; 2BNH.
DR Genew; HGNC:16400; CIAS1.
DR MIM; 606416; -
DR MIM; 120100; -
DR MIM; 191900; -
DR MIM; 607115; -
DR GO; GO:0006917; P:induction of apoptosis; NAS.
DR GO; GO:0006954; P:inflammatory response; IMP.
DR GO; GO:0007165; P:signal transduction; NAS.
DR InterPro; IPR001611; LRR_RNinh.
DR InterPro; IPR007091; LRR_RNinh_sub.
DR InterPro; IPR003590; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN.
DR Pfam; PF00560; LRR; 6.
DR Pfam; PF05729; NACHT; 1.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PRINTS; PRO0019; LEURICHRPT.
DR SMART; SM00368; LRR_RI; 3.
DR PROSITE; PS00824; DAPIN; 1.
DR PROSITE; PS00837; NACHT; 1.
KW Alternative splicing; Apoptosis; Deafness; Disease mutation;
KW Leucine-rich repeat; Repeat.
FT DOMAIN 1 91 DAPIN.
FT DOMAIN 218 534 NACHT.
FT REPEAT 738 761 LRR 1.
FT REPEAT 795 818 LRR 2.
FT REPEAT 852 875 LRR 3.
FT REPEAT 881 904 LRR 4.
FT REPEAT 909 937 LRR 5.
FT REPEAT 939 961 LRR 6.
FT REPEAT 966 989 LRR 7.
FT DOMAIN 688 695 Poly-Glu.
FT VARSPLIC 718 1034 Missing (in isoform 3).
FT VARSPLIC 719 775 Missing (in isoform 1).
FT VARSPLIC 834 890 Missing (in isoform 1).
FT VARSPLIC 198 198 Missing (in isoform 1).
FT VARSPLIC 260 260 Missing (in isoform 1).
FT VARSPLIC 303 303 Missing (in isoform 1).
FT VARSPLIC 305 305 Missing (in isoform 1).
FT VARSPLIC 309 309 Missing (in isoform 1).

```

FT VARIANT 348 348 /FTid=VAR_014106.
FT T -> M (in MMS).
FT /FTid=VAR_014366.
FT VARIANT 352 352 A -> V (in MMS).
FT /FTid=VAR_013228.
FT VARIANT 358 358 H -> R (in CINCA).
FT /FTid=VAR_014367.
FT VARIANT 436 436 T -> N (in CINCA).
FT /FTid=VAR_014368.

Query Match 43.8%; Score 2396; DB 1; Length 1034;
Best Local Similarity 46.1%; Pred. No. 1.1e-161;
Matches 495; Conservative 175; Mismatches 303; Indels 100; Gaps 10;

QY 11 CRISTYLEEELAEVLEKFKLYL-GTATELGEGKIPWGSMEKAGPLEMAQLLTHFGPEEA 69
DB 6 CKLARYLEDLEDVDLKKFKHLEDPYPPQKGIPLPRGQTEKADHVDLATLMDFNGEKA 65
QY 70 WRLALSTFERINRKNLWEGORED----- 93
DB 66 WAMAVWTFAPAAINRRDLYEKARDEPKWGSNARVSNPTVICQEDSIEEEMWGLLEYLSRI 125
QY 94 ----LVDPDPOETVYRRFRLEMDERNARIGECVNLSHRYTRILLVKEHNSPMQVQOOL 149
DB 126 SICWKKDYRKRYKRYRSFQIEDNRNARGESVSLNKRYTRLRIKEHRSQOEREQEL 185
QY 150 LDTGRGHARTVGHQASPIKETIPEPDEERPEPPTVVMQAGAGIKSMIAHKVMDLWDAD 209
DB 186 LAIGK--TKCESPVSPKIMELLFDPDEHSEPVHTVVFQAGAGIKTILARKXMDLWDAS 243
QY 210 GKLFQGRFDYLYINCREMNSQATECSMDQLIFSCWPEPSAPIQELIRVPERLLFIIDGF 269
DB 244 GTLYQDRFDYLYITHCREVS-LVTRQSLGDLINSCCPDPNPIHKIVRKPSRIILFLMDGF 302
QY 270 DELKPSFHDQGPWCLWEKRPTELALNSLRKLAPELSLITPTALEKHLRLLEH 329
DB 303 DELQAFDEHIGPLCTDQKAERGDIILSLRKLLEPEASLLITTPVALEKLQHLDDH 362
QY 330 PRHVEILGFSEAEKKEYFYKFNHAEQAGQVFNVRDNEPLFTMCFVPLVCVVCTCLOQ 389
DB 363 PRHVEILGFSEAEKKEYFYKFNHAEQAGQVFNVRDNEPLFTMCFVPLVCVVCTCLOQ 422
QY 390 QLEGGLLROTSRTTTHVYMLYLLSMQPKGAPRLQPPNQKGLCSLAADGKLNQKILF 449
DB 423 QMESGKSLAOTSRTTTHVYMLYLLSMQPKGAPRLQPPNQKGLCSLAADGKLNQKILF 482
QY 450 BEODLRKHGIDGEDVSAFLNWNIFOKDINCERYYSFTHLSFQFEFAAMYILDEGGAG 509
DB 483 EESDLRHHGQKADVSAFLNWNIFOKDINCERYYSFTHLSFQFEFAAMYILDEGGAG 542
QY 510 -----PDQVTRLLTYAFSERSFLALTSRFLFGLNNEETRSHEKSLCKWVSPHI 560
DB 543 NVFGSRLLKPSRDVTVLLENYGRFEKGYLIFVVRFLFGLNVQERTSVLEKLSCKLSQOI 602
QY 561 KMOLLQWIKQASGDSTLQOGLSFPSCYIYETQEEFIQOALSHFOVIVVSNIAKMEH 620
DB 603 RLELLKWEIVKAKAKKIQIOPSQLELYFYLYEQEEDFQVAMDYPPKIEI-NLSTRMDH 661
QY 621 MVSSFCCLKRCSAQVHLHYNTYASDGEDRARCAGATHLLVQLRPBTVLLDAYSEHLA 680
DB 662 MVSSFCLENCHRVESLSL-GFLNMPKXEBEEKEGRHLDVQ----- 703
QY 681 AALCTNPNLTSLRYNALSGRVKLLCQGLRHPNCKQLNLRKCRISSSACEDLSAAL 740
DB 704 ---CVLPs-----SSHAACSHG-----LVNSHLTSSFCRGLFSVL 735
QY 741 IANKNLTRMDLSNGVGFPGWMLLCEGLRHPQCRLOMIQRLKQLESGAGCAQEMVLGTN 800
DB 736 STQSRLTELDSNLSLGDPMGVLCETLQHPGCMIRRLWLGRCLSHECCFDLSVLSSN 795
QY 801 PHLVELDLTGNALEDGLRLLCQGLRHPVCLRLTLWLKICRLTAACADELASTLSVNQSL 860
DB 796 QKLVELDLSNALGDGFGIRLLCVGLKHLNCLNKLVLVSCCLTSACQDLASVLSTSHSL 855

```

```

QY 861 REIDLSLNEIGDGLVLLLCGLRHPTCKLQTLRLIGICRLGSAACEGLSVVLQANHNLREL 920
DB 856 TRLYVGENAGLDSGVAILCEKAKNPQCNQKLGVLNSGLTSVCCSALSLSVLSTQNLTHTL 915
QY 921 DLSFNLDGDLWLLAEGLOHPACRLQKLDSCGLTAKACENLYFTLGINOTITDLYLT 980
DB 916 YLRGNTLGDGKIGKLLCEGLLHPDKLQVLELDNCLNLTSHCCWDLSTLLTSQSRLKLSLG 975
QY 981 NNALGDTGVRLLCKRLSHPGCKRLVRLWFGMDLNKMTSHRLAALRVTKPVLDI 1033
DB 976 NNDLGDGLVWFMFEVLKQOQSCLLQNLGLSEMYFNYSKSALETLOEKEPCLTV 1028

RESULT 4
CIS1_MOUSE
ID CIS1_MOUSE STANDARD; PRT; 1033 AA.
AC QBR4B8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Cold autoinflammatory syndrome 1 protein homolog (PYRIN-containing
DE APAF1-like protein 1) (Mast cell maturation inducible protein 1).
DE Names-Cisai; Synonyms=Mnigi, NALF3, Pypafi;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=BALB/cJ;
RA Kikuchi-Yanoshita R., Koga K., Taketomi Y., Sugiki T., Saito T.,
RA Ishii S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,
RA Chang H.-W., Sawada M., Inagaki N., Nagai H., Murakami M., Kudo I.;
RT Identification of inducible genes during in vitro maturation of mouse
RT bone marrow-derived mast cells to connective tissue-type mast cells.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: May function as a potential inducer of apoptosis.
CC Interacts selectively with apoptosis-associated specklike protein
CC containing a CARD domain (ASC). This complex may function as an
CC upstream activator of NF-kappa-B signaling (By similarity).
CC -1- SIMILARITY: Contains 1 DAPIN domain. (LRR) repeats.
CC -1- SIMILARITY: Contains 7 leucine-rich
CC -1- SIMILARITY: Contains 1 NACHT domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DB EMBL; AF486632; AAL90874.1; -
DB HSSP; P10775; 2BNH.
DB MGD; MGI:2653833; C1a1.
DB InterPro; IPR001611; LRR.
DB InterPro; IPR007091; LRR_Nrnh.
DB InterPro; IPR007111; NACHT_NTPase.
DB InterPro; IPR004020; PAAD_DAPIN.
DB Pfam; PF00560; LRR; 6.
DB Pfam; PF05729; NACHT; 1.
DB Pfam; PF02758; PAAD_DAPIN; 1.
DB PRINTS; PR00019; LEURICHPPT.
DB PROSITE; PS50824; DAPIN; 1.
DB PROSITE; PS50837; NACHT; 1.
DB Apoptosis; Leucine-rich repeat; Repeat.
DB DOMAIN 1 91
DB DAPIN.
DB FT DOMAIN 216 532
DB FT REPEAT 737 760
DB FT REPEAT 794 817
DB FT REPEAT 851 874
DB FT REPEAT 880 903

```

QY 921 DLSNDLGDWGLMLLAEGLOHPACRQLKWLDSOGLTAKACENUYFTLGINGNQTILDTLYLT 980

Db 915 YLRNALGDDTGRLLECEGLHDPCKQLMDLNCSLTSHSCWNLTSTTLTHNHSLRKINLG 974

QY 981 NNAUGDGTGVRLLCRLSRHSFGCKLRVLVLFMGDLNKWTHSRLAALRRVTKPYLDI 1033

Db 975 NNDLGDLCVVTLCEVLKQQCGCLQSLSQGEMYLNRRTKRAELAQEKPELTI 1027

RESULT 5

Q6TG35 PRELIMINARY; PRT; 977 AA.

AC Q6TG35;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Cryopyrin deletion 4 isoform.

GN Name=CIAS1;

OC Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI_TaxID=9606;

RN [1]_SEQUENCE FROM N.A.

RP PubMed:1466282;

RX O'Connor W. Jr., Harton J A., Zhu X., Linhoff M.W., Ting J.P.Y.;

RT "CIAS1/Cryopyrin/PYPAFI/NALP3/ CATERPILLER 1.1 Is an Inducible

RT Inflammatory Mediator with NF-kappaB Suppressive Properties.";

RL J. Immunol. 171:6329-6333(2003).

DR EMBL; AY422168; AAC98889.1; -.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR007091; LRR_Rninh.

DR InterPro; IPR003590; LRR_Rninh.sub.

DR InterPro; IPR007111; NACT_NTPase.

DR InterPro; IPR004020; PAAD_DAPIN.

DR Pfam; PF00560; LRR_1; 4.

DR Pfam; PF02729; NACT; 1.

DR Pfam; PF02758; PAAD_DAPIN; 1.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00368; LRR_R1; 7.

DR PROSITE; PS50824; DAPIN; 1.

DR PROSITE; PS50837; NACT; 1.

SQ SEQUENCE 977 AA; 111624 MW; F17E49B93563A7184 CRC64;

Query Match 41.2%; Score 2253.5; DB 2; Length 977;

Best Local Similarity 44.1%; Pred. No. 1.4e-151;

Matches 473; Conservative 164; Mismatches 279; Indels 157; Gaps 11

QY 11 CRUSTYLEAVEALKKKLYL-GTATELGBGKIIPWCSMEKGAPLEMAQILLITFHGPESA 69

Db 6 CKLARVLEDELVDLKKFMHLEDYPPOKGCIPLPQGTETKAHDVLDLATLMDFNGEEKA 65

QY 70 WRLLATFFIRNKOLWERGORE----- 93

Db 66 WANAVIFAANRRDLRYEKARDEPKWGSDNARSNTVICQEDSIIEEWMLLEYLSRI 125

QY 94 -----LVRDQETTYDVYVRKFLMEDRNARGECVNLSHRVTRLLLVKEHSPMQVQQOL 149

Db 126 SICWKMKDYRKRYKYVSRSFQCIEDRNARGESVSNKRYTRLRLIKEHRSQOEREQEL 195

QY 150 LDTCGRGHARTVGHQASPIKETLFPDBEPPEPRTVMQGAAGIGKSMIAHKVMLDWAD 209

Db 186 LATGK--TKCESPVSPINKELLFPDDESEPHVTVFQGAAGIKGITLARKOMLDWAS 243

QY 210 GKLFQGFDFYLVFNCREMNNQATSCSQODLIIFSCWPSPSAploELIRVPERLLFIIDGF 269

Db 244 GTLYQDRFDYLVFIHCREYS-LVTQRSGLDILMSCPDNPPIHKIVKPSRILFLMDGF 302

QY 270 DELKPSFHDPPQGWPLCWEEKPTEKPTALLNSLTRKKLLPELSLLITTPTALEKLHRLLEH 329

Db 303 DELQCAFDEHGIPLCDWDQAERGDIILLSLRKKLLPEASILLITTPVALEKLQHLLDH 362

QY 330 PRHVIELGFSEARKEYFYKHFAAQOVNYVRDNPELFTWCFLVCVWVCTCLOQ 394

| | | | | | |
|----|------------------------|---|---------------------|----------------------|--------------|
| FT | REPEAT | 908 | 931 | LRR 5. | |
| FT | REPEAT | 937 | 964 | LRR 6. | |
| FT | REPEAT | 965 | 988 | LRR 7. | |
| FT | SEQUENCE | 1033 | AA, 118274 | MM, 592469096GB12117 | CRC64; |
| QY | Best Match | 43.1% | Score 2359.5; | DB 1; | Length 1033; |
| QY | Query Local Similarity | 46.6% | Pred. No. 4.2e-159; | Indels 101; | Gaps 13; |
| QY | Matches 500; | Conservative 165; | Mismatches 307; | | |
| QY | 11 | CRLSLLEELAEVLEKFKLYL-GTATELGEKIPWGSXAGGLEMAQLLITHOPERA | 69 | | |
| QY | 6 | CKLAQYLEDLEDVLLKFKHLEDYPPKGCIPVRGQMEKADHLDTALMIDFNGEKA | 65 | | |
| QY | 70 | WRLALSTFERINKDWERQOR-----EDLV----- | 95 | | |
| QY | 66 | WAMAVIFAAINRDLWEKAKQDPENWDCTTSHSSVMVCQDLSLEEWGMLGYLSRISI | 125 | | |
| QY | 96 | ----RDQETRYDVARKFELMEDNARLGEVCNLSHRYTRLLLVKXHSNPMQVQOQLD | 151 | | |
| QY | 126 | CKKKDKYCKMYRRHVSRSRFSIKDNARLGEVDLSNRYTQLVLEKHPKQREHELLT | 185 | | |
| QY | 152 | TGRGHARTVGHQASPIKIEFLPPDEBERPEPRTVVMQGAAGIGKSMIAHKWMLWDAGK | 211 | | |
| QY | 186 | IGRTQWR--DSPMSSLKLELLEPDEGHSEPHVTTFVQGAAGIGKTLARKIMLDWALGK | 243 | | |
| QY | 212 | LFGREFDYLVINCENWQNSATCSMODLIFSCWPEPSAPLOELIRVPERLLFTIDGFE | 271 | | |
| QY | 244 | LFKQKFDYLFTHCREVS-LRTPRSADIIVSCWDPNPPVCKILRKESRILFLMDGFE | 302 | | |
| QY | 272 | LKPSFHPDQPGWCLWCBEKPTTELLNSLRKKLPELSLITTPPTALEKHLRLHSHPR | 331 | | |
| QY | 303 | LQAGFDEHIGVCTQKAVRGDILLSSLRKKLLPKASLLITTPPALEKHLQHLLOHPR | 362 | | |
| QY | 332 | HVEILGFSEAKRKYFYKYHNAEOAGOVNVRDNEPLFTWCFVPLVCVWVCTCLOQOL | 391 | | |
| QY | 363 | HVEILGFSEAKRKYFYKYSNELQAREAFRLQENEVLTWCFPLVCWVCTGLKQOM | 422 | | |
| QY | 392 | EGGGLRLQTSRTTAVTVMLYLLSLMQPKGAPLQPPNORGLCSLAADGLWNOKILPEE | 451 | | |
| QY | 423 | ETGKSLAQTSKTTTAVTVVFLSLLQSRGGIEHLFSYDLQGLCSLAADGTVNOKILPEE | 482 | | |
| QY | 452 | QDLRKHGLDGEDVSAFLNMNLFKQDINCERYYSFIHLSQEFPAAMYITLDE----- | 503 | | |
| QY | 483 | CDLRKHGLQKTDVSAFLRMNVFQKEVDCERFYFSHMTQEFPAAMYILLEEABGTVR | 542 | | |
| QY | 504 | -GEGGAGP--DQDVTRLITAYAFERSFLATGRPLFGLLNBTSHLEKSLCWKVSPHI | 560 | | |
| QY | 543 | KFGPGCSDLLNRDVKVLLNENYKPEKGYLIFVVRFTFGLVNBQRTSYLEKKLSCKISQOV | 602 | | |
| QY | 561 | KMDLLOWTSKAQSDGSTLQOGLSEFFSCLYETQEEFIQALSHPOVIVWSNIASKWEH | 620 | | |
| QY | 603 | RLELLKWIEVKAKKLQWQPSQLEFVCLYEMQEDDFVQSAMHHPKTEI-NLSTRMDH | 661 | | |
| QY | 621 | MYSSFCLCKCRSAQVHLHYGATYSADGEDRARCAGAHITLLVLRPERTVLLDAYSEHLA | 680 | | |
| QY | 662 | VVSSFICKNCHRVTLSL-GFFHNSPKEEBERGG-----RP-----LD | 700 | | |
| QY | 681 | AALCTPNPILIELSVYNALGRGVKLLCQGLRHPNCKLQNLKRCRISISSACEDLSAAL | 740 | | |
| QY | 701 | QVQCVPFD-----THVAC---SSRLVNCLTSFSGRGLFSSL | 734 | | |
| QY | 741 | IANKNLTRMDLNGVGGPFGMMLLCEGLRHQPQCRLOMQIRKQLQESGACQEMASVLGPN | 800 | | |
| QY | 735 | STRSRITELDSNTLGDGFMEVLCEALQHPGCNIQRLGRGCLSHQCCFISSVLSSS | 794 | | |
| QY | 801 | PHLVELDLTGNALDEDLGLRLLCQGLRHPVCRILRTILWLKICELTAACDELASTVNSQSL | 860 | | |
| QY | 795 | QKVLVDLSDNALGDFGIFLLCVGLKHLNQLKWLVSCLTSACQDIAUVLSSNHEL | 854 | | |
| QY | 861 | RELDLSINELGDLGVLLLCGLRHPTCKLQTLRLGICRLGSAACEGLSVVLQAHNLREL | 920 | | |
| QY | 855 | TRLYIGENALGDSGVQVLEKMKDQCNLQKGLVNSGLTSCCSALTSVLKTNQNFTHL | 914 | | |

Db 363 PRHVEILGFSAKREYFFKFSDEAQAARAFSLQIENEVLFTWCFPLVVCWIVCTGLKQ 422
Qy 390 QLEGGGLLRQTSRTTAYVLMYLLSLMOPKPGAPRLQPPNORGLCSLAADGLNWKILF 449
Db 423 QMESGKSLAQTSKTTAYVFLSSLLQPGSGQEHGLCAHLWGLCSLAADGIWNKILF 482
Qy 450 EQDLRKHGLDGEDVSAFLNINIFQKINDICRYYSFIHLSQEPFFAAMYIILDEGEAG 509
Db 483 EESDLRNLGLOKADVSAPLNLNLPQKEVCEKFSFIHMTQEPFFAAMYIILDEGEAG 542
Qy 510 -----PDQDTRILTEAFERSFLATSLTFLGLNEETSHLEKSLCWKVSPI 560
Db 543 NVPSRLKLSRDVTVLLENYKPEKGYLFFVFRFLGVLNQERTSYLEKSLCKISQOI 602
Qy 561 KMDLLQWTSQAQSGDSTLQOQSLEFFSCYIEQEIEFIQALSHFQVIVVSNIASKWEH 620
Db 603 RLELLKWEVKAKAKKQIQPSQLEFLCYLMEDEDFVQAMDYFPKIEI-NLSTRMDH 661
Qy 621 MVSSFCILKRCRSAOVHLHYGATYGADGEDRARCAGAHLLVQLRPRTVLLDAYSEHLA 680
Db 662 MVSSFCIENCHRVESLSL-GPLHNPKEEBEERKGRHLDVQ----- 703
Qy 681 AALCTNPNLIELSYRNALSGRYKLLCOGLRHPNCKLQNLRLKRCRISSACEDLSAAL 740
Db 704 ---CVLPS-----SSHAAC----- 714
Qy 741 IANKNLTRMDLSGNGVGPFGMMLCEGLRHPQCRLOLQRLKCOLSGAGCAQEMASVLTGN 800
Db 715 -----SGLG-----RCGLSHECCFDLSVLSSN 738
Qy 801 PHVELDLTGNALEDGLRLCCOGLRHPVCRRLTLWLKICRLTAACADELASTLVNQSL 860
Db 739 QKVELDLSNALGDFGIRLLCVGLKLLCNLKLWLVSCCLTSACQDLASVLSTSHSL 798
Qy 861 RELDLSNELGDLGVLLCEGLRHPTCKLQTLRLGICRLGSAACEGLSVLQAHNNREL 920
Db 799 TRLYVGENALGDSGVAILCEKAKNPQCNLQKGLVNSGLTSVCCSALSVLSTNQNLTHL 858
Qy 921 DLSFNDLGDWGLWLLAELGLOHPACRLQKLDSDGLTAKACENLYFTLGINOTLTLTLYLT 980
Db 859 YLRGNTLGDGKILKCEGLLHPDCKLQVLELDNCLNLSHCCWDLSTLTSQSQRKLSLG 918
Qy 981 NNALGDTGVRLLCRLSHPGCKRLVWLFGMDLNKMTSHSLAALRVTKPYLDI 1033
Db 919 NNDLGLGVMMFCEVLKQOSCLLQNLGLSEMYFNYETKSALETQEEKPELTV 971

RESULT 6

NA14 HUMAN STANDARD; PRT; 1093 AA.
ID NA14 HUMAN 32.9%; Score 1799; DB 1; Length 1093;
AC Q86W24; Q7RTR6;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE NACHT-, LRR- and PYD-containing protein 14 (Nucleotide-binding
oligomerization domain protein 5).
GN Name=NA1P14; Synonyms=NOD5;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22451042; PubMed=12563287; DOI=10.1038/nrml1019;
RA Tschopp J., Martinon F., Burns K.;
RT "NALP5: a novel protein family involved in inflammation.";
RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22651535; PubMed=12766759; DOI=10.1038/nrml1086;
RA Inohara N., Nunez G.;
RT "NODs: intracellular proteins involved in inflammation and

RT apoptosis.";
RL Nat. Rev. Immunol. 3:371-382(2003).
CC -!- FUNCTION: Involved in inflammation (Potential).
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 11 leucine-rich (LRR) repeats.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AY154469; AA018165.1; -;
DR EMBL; BK001107; DAA01240.1; -;
DR HSPSP; P10775; 2BNH.
DR Genew; HGNC:22939; NALP14.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN.
DR Pfam; PF00560; LRR; 8.
DR Pfam; PF05729; NACHT; 1.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR PROSITE; PS00824; DAPIN; 1.
DR PROSITE; PS00837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 1 97 DAPIN.
FT REPEAT 177 499 NACHT.
FT REPEAT 525 550 LRR 1.
FT REPEAT 672 695 LRR 2.
FT REPEAT 733 756 LRR 3.
FT REPEAT 757 780 LRR 4.
FT REPEAT 785 807 LRR 5.
FT REPEAT 842 869 LRR 6.
FT REPEAT 871 894 LRR 7.
FT REPEAT 899 926 LRR 8.
FT REPEAT 956 979 LRR 9.
FT REPEAT 985 1008 LRR 10.
FT REPEAT 1013 1037 LRR 11.
SQ SEQUENCE 1093 AA; 124732 MW; 124EEACE22A11D6F CRC64;
Query Match 32.9%; Score 1799; DB 1; Length 1093;
Best Local Similarity 37.4%; Pred. No. 4e-119;
Matches 413; Conservative 177; Mismatches 388; Indels 126; Gaps 15;
Qy 13 LSTYLELEAVELEKFKLYLGTATLGEKIPWGSMEKAGPLEMAQLLIHFGPPEAWRL 72
Db 15 LLLYLEELNKEELNFTKLFKETWEPHGLTPNNEVKARREDLANLKKYYPGEKASV 74
Qy 73 ALSTFERINRKLWERGORE-----DLVRDPQET-----YRDYVRRKF 110
Db 75 SLKIFGKMLKOLCERAKEINNSAQTIGPDDAKAGETQDQBAVLGDGTEYRNRIKEKF 134
Qy 111 RLMEDNRALGECVNLSHRYTRLLLVKHSNPNQVQQQLLDTGRGHARTVGHQASPIKE 170
Db 135 CITWKKSLAGKPEDFHGG-----TAKDRKL-----LE 163
Qy 171 TLFEPEDEERPEPRTPVVMQGAAGIKGSMALAHKVMLOWADGKLFQGRDYLYPINCENMQ 230
Db 164 HLFVDVKTGAQFIQVVLQGAAGVGTLLVRKAMLDWAGSLVQQRKYVFLNGRINQ 223
Qy 231 SATECSNQDLIFSCWPSPSAPLOELIRVPERLLFIIDGFDLKPSPHDPQGMCLCWEK 290
Db 224 -LKERSFAQLISKDWPSTEGPIEIMVQPSLLFIIDSFDLNFAPFEPFALCEDWTQE 282
Qy 291 RPTPELLNSLRKLLPELSELTITPTALEKLRHLEHPRHVEILGFSAEKREYFYKY 350
Db 283 HPVSFLMSSLLRKVMLEASLLVTRTLTTSKRLKQLLKNHHYVELLGMSEDAEETIYQF 342

-!- MISCELLANEOUS: The sequence shown here is derived from an entry /GorBank/0001 third party annotation (TPA) entry.

| | | | |
|----------|---|--|--------------|
| QY | 351 | PHNBAQAGGVNFYVRDNPBLFTMCVPLVWVWCTCLOQOEGGGLLRQTSRTTAVYML | 410 |
| Db | 343 | PEDKWAQVFSGLKSNEMLFMCQVPLVWCAAACTCLAQMEKGGDVTLTCOTTITALFTC | 402 |
| QY | 411 | YLLSLMQP-KPGAPRIQPPNCO---RGLCSLAADGLWQKILFEQDQKRLKHGLDGEDVSA | 466 |
| Db | 403 | YISSUFTPVDGGSPL---PNAQALRLCOVAAKGIWTWYVYFRENLRRLGLTQSDVSS | 459 |
| QY | 467 | FLNNMIFQKDINCERYYSFIHLSPOEFFAAMYIL-----DEGGAGAPDQDTRILLTEYA | 522 |
| Db | 460 | FMSDNIIQKDAEYENCYVFTLHVQEFFAAMFYMLKGSWEAGNPSQPEPDLKSLQSTS | 519 |
| QY | 523 | PERSFIATSRPLFGLINEETRSLEKLSLWKVSPHISKWIDLLQWIOSQAQDGSGLTQOG | 582 |
| Db | 520 | YKD-PHLTQMKCFGLLNEDRVKQURTFNCWMSLKISKLLQCEVGLGNSDYSPSQLG | 578 |
| QY | 583 | SLEFFSCLEYIOEBEETIOQALSHFOVIVVSNISAKMEHVVSPFCLKRCRQAQVILHLYGA- | 641 |
| Db | 579 | FLELPHCLVETQDAFISQAMRCFPKVAI-NICEKHLLVSSPCLKCHCCLRTIRLSVTV | 637 |
| QY | 642 | -----TYSADGEDRARC-----SAGAHTLVQLRPERTVLLDAYSEHL----- | 679 |
| Db | 638 | VFEKYLKTSPLPTNTWDGDRITHCWQDLCEVLHT-----NEHLRELDIYHNSLNXSAMN | 691 |
| QY | 680 | -----AAALCTNPNIIELSLYRNALGSRGVKLLCQ | 709 |
| Db | 692 | ILHHELHPNCKLQKLLKFTFPDGCQDIDTSLIHNKMLHMLDLKGSIDGNGVKSLCE | 751 |
| QY | 710 | GLRHPNCKLQNLRLKCRISACEDLSAALIANKNLTRMDLSNGVGPFGMMLLCEGLR | 769 |
| Db | 752 | ALKHPECKLQTLKLESCNLTAVFCNLINSNALIRSQSLFPLNLTNNLLDGVQVLLCEALR | 811 |
| QY | 770 | HPQCLQMIOLRLKQLESAGACQEMASVLGTNPPLVELDTGNALEDLGLFLLCOGLRHPV | 829 |
| Db | 812 | HPKYLRLSLESGLTEAGEVLSLALISNKRLLTHLCLADNVLDGQGVKLSMDALQHAQ | 871 |
| QY | 830 | CRRLTLWKTCRLTAACDEBLASTLSVNSQSLRELDLSNELGDLGVLLLCEGLRHPTCKL | 889 |
| Db | 872 | CTLKSLVLRCHFTSLSEYLSLTLNKSLLTHLDLGSNLQDQNGVKLLCDVFRHPSCNL | 931 |
| QY | 890 | QTLRLGICRLGSAACEGLSVVLQAHNHLRELDLSFNDLGDWGLWLLAEGI-OHPACRLQKL | 949 |
| Db | 932 | QDLMLGCVLTNACCLDLASVILANNPNRLSLDLGNNDLQDDGVKILCDALRYPNCNTQRL | 991 |
| QY | 950 | WLDSCGLTAKACENLYFTLGINQTLTDLIYTNNAALGDTGVRLLCRKLHSHPGCKLRVLWL | 1009 |
| Db | 992 | GLEVCGLTSLCCQDSSALICNKRLLIKONLTQNTLGVGIVKLYKLSFKPKRLQVLGLC | 1051 |
| QY | 1010 | GMDLNKMTSHSLAALRVTKPYLDI | 1033 |
| Db | 1052 | KEAFDEBAQKLEAVGVSNPHLII | 1075 |
| RESULT 7 | | | |
| Q7RTR0 | | PRELIMINARY; | PRT; 991 AA. |
| AC | Q7RTR0; | | |
| DT | 01-WAR-2004 (TREMELrel. 26, Created) | | |
| DT | 01-WAR-2004 (TREMELrel. 26, Last sequence update) | | |
| DT | 01-WAR-2004 (TREMELrel. 26, Last annotation update) | | |
| DE | NOD6. | | |
| OS | Name=NOD6; | | |
| GN | Homo sapiens (Human). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| OX | NCBI_TaxID=9606; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RX | MEDLINE=22651535; PubMed=12766759; | | |
| RA | Inohara N., Nunez G.; | | |
| RT | "NOD5: Intracellular proteins involved in inflammation and apoptosis."; | | |
| BL | Nat. Rev. Immunol. 3:371-382(2003). | | |


```
QY 782 KCQLESGAQEMASVLTGNPHLVLDLTGNALDGLRLLLCOGLRHPVRLRTLWLKICR 841
DB 723 KCDISSEVCEDIASVLACNSKLKHLSLVENPLRDEGMTLLCEALKHSHCALERLMLMYCC 782
QY 842 LTAACDELASTLSVNQSLRELDLSLNEGLDGLVLLICEGLRHPTCKLOTIRLIGICRLGS 901
DB 783 LTSVSCDSISEVLLCSLSLDDLSGNALEDNGVASLCAALKHPGCCSIRELWLMGCFLLTS 842
QY 902 AACBGLSVVLQANHLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKLWLDSCGLTAKAC 961
DB 843 DSCKDIAAFLVLCNGKLTLLGHNIEGDTGVRQLCAALQHPCKLECLGLQTCPIRACC 902
QY 962 ENLYFTLGINOTLTDLYLTNNALGDTGVRLLCKRLSHPGCKRLVLMFGMDLNRKTHSRL 1021
DB 903 DDIAAALIACTLRSLNLDWIALDADAVVVLCEALSHDPDCALQMLGLHKSFGDETKIL 962
QY 1022 AALRVTKPYLDI 1033
DB 963 MSVEKIPHLTI 974

RESULT 8
Q86W27 PRELIMINARY; PRT; 986 AA.
AC Q86W27;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE NALP9
GN Name=NALP9;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22451042; PubMed=12563287;
RX Tschopp J., Martinon F., Burns K.;
RT "NALPs: a novel protein family involved in inflammation.";
RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
DR EMBL; AY154464; AA018160.1; -.
DR HSSP; P10775; 2BNH.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN.
DR InterPro; IPR001865; Ribosomal_S2.
DR Pfam; PF00560; LRR_1; 4.
DR Pfam; PF05729; NACHT; 1.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; UNKNOWN_1.
DR PROSITE; PS00824; DAPIN; 1.
DR PROSITE; PS00837; NACHT; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; UNKNOWN_1.
SQ SEQUENCE 986 AA; 112592 MW; B4248B46F70413B2 CRC64;
```

```
Query Match 27.5%; Score 1505.5; DB 2; Length 986;
Best Local Similarity 34.2%; Pred. No. 2.9e-98;
Matches 344; Conservative 193; Mismatches 383; Indels 85; Gaps 18;

QY 16 YLELEAVELEKFKLYLGTATPELGEK-IPGSMKEKAGPLEMAQLLITHFGEPAWRLLAL 74
DB 14 YLKLKEEFKFKELKQLPELKPDPWAEKKASKEDVAKLLDKHVPKGQWVTL 73
QY 75 STFERINRKLWEGQREDLVDPQETRYDYVRKFRIMEDNRNARLGEVNLSHRYTRL 134
DB 74 NLFLQINRKLWLWKAQEE--MRNKLNPYRKHKMTFQLWEKET---CLHVPHEFYKET 127
QY 135 LVKESHSPMQVQQLDGTGRGHARTVGHQASPIKIELTFEPDERPPTVTVQGAAGI 194
DB 128 MKNEY-----KELNDAYTAARR-----HTVVLGEGDGI 156
```

```
QY 195 GKSLAHKVMLDWADGKLFQGRFDYFYINCREMNOGATECSMODLIFSWPEPSAPLOQ 254
DB 157 GKTTLRLKVMLDWAEGLNWKDRFTTFVFLNVCENWGIA-ETSLLELLSRDWPESSEKIED 215
QY 255 LIRVPERLLFIIDGFD-----ELKPSFHPDQPGWCLCWEKRTTELLLSLIRKLLPE 308
DB 216 IFSQPERILFIMDGFEQLKFNLOKADLSD-----WRQRPMPILSLLOKQMLPE 268
QY 309 LSLITTRPALEKHLLEHPRHVEITLGFSEAEKKEYFYKYPHNAQAGOVFNVRDNE 368
DB 269 SLLIALGKLMQKHYPMLRHPKLIKULGFSESEKKSIFYFGEKSKALKVFNVRDNG 328
QY 369 PLFTMCVPLVWVCTCTCQOQLEGGLRQTSRTTTAVYMLYLLSLMKPKPGAPRLOPP 428
DB 329 PLFTLCHNPFTCMVCTCVQRLGERGDLINSQNTTYLYASFLTVF--KAGSQSPPK 386
QY 429 PNQ---RGLCSLAADGLWNOKILPEBODLRKHGLDGEDVSAFLNMNIFQKDINCERYSP 485
DB 387 VNRARLKSALAAEGITTYTFVSHGDLRRNGLSESEGVMMVGMRLLRGGDC---PAF 443
QY 486 IHLSPQEFPAAMYILDEGEGGAGPD-QDVTRLLITEYAFSERSFLALTSRFLGLLNEET 544
DB 444 MHLCIQEFCAAMFYLLKRPDDNPALGSITQLVRASVVOPTLLTVQVGFMPGISTEEI 503
QY 545 RSHLEKSLCWKSPHIKMDLLQIQSKAQSDGSTLOQGSLEFFSCLYEIOEEEFIOQALS 604
DB 504 VSMLETSFGFPLSKDLAQEITQCLESLSQCEADREATAFOELFGLFETQKEFVTKVMN 563
QY 605 HFQ--VIVVSNIASKMEHMY--SSFCLKRCRSQAQVHLHYGATYSADGEDRARCAGAHULL 661
DB 564 FPEEVFIYIGNI---EHLVIASFCLKHCQHLTLTRMCVENIFPDD---SGCISDYNEKL 616
QY 662 VOLRPERTVLLDAYSEHLAALCTNPMLIELSVRNALGSRGVKLLCOGLRHPNCKLQNL 721
DB 617 V-----YRELCSMFTTNKNFQILDMENTSLLDDPSLAILCKALQAPVCKLRKL 664
QY 722 RLKRCRISSACEDLSAALIANKNLITRMDLSGNGVGFPGMMMLCEGLRHPCRLQMLQLR 781
DB 665 IFTSVYFGHDS--ELFKAVLHNPHLKLLSLVGTSLQSQSDIRHLCETLKHPWCKIEELILG 722
QY 782 KCQLESGAQEMASVLTGNPHLVLDLTGNALDGLRLLLCOGLRHPVRLRTLWLKICR 841
DB 723 KCDISSEVCEDIASVLACNSKLKHLSLVENPLRDEGMTLLCEALKHSHCALERLMLMYCC 782
QY 842 LTAACDELASTLSVNQSLRELDLSLNEGLDGLVLLICEGLRHPTCKLOTIRLIGICRLGS 901
DB 783 LTSVSCDSISEVLLCSLSLDDLSGNALEDNGVASLCAALKHPGCCSIRELWLMGCFLLTS 842
QY 902 AACBGLSVVLQANHLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKLWLDSCGLTAKAC 961
DB 843 DSCKDIAAFLVLCNGKLTLLGHNIEGDTGVRQLCAALQHPCKLECLGLQTCPIRACC 902
QY 962 ENLYFTLGINOTLTDLYLTNNALGDTGVRLLCKRLSHPGCKRLVLMFGMDLNRKTHSRL 1006
DB 903 DDIAAALIACTLRSLNLDWIALDADAVVVLCEALSHDPDCALQML 947

RESULT 9
Q86U61 PRELIMINARY; PRT; 506 AA.
AC Q86U61;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Cryopyrin (Fragment).
GN Name=CIAS1;
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RX PubMed=15302403; DOI=10.1016/j.gene.2004.05.002;
RA Anderson J.P., Mueller J.L., Rosengren S., Boyle D.L., Schaner P.,
RA Cannon S.B., Goodyear C.S., Hoffman H.M.;
RT "Structural, expression, and evolutionary analysis of mouse CIAS1.";
RL Gene 338:25-34(2004).
DR EMBL; AY338196; AAR03567.1; -.
DR InterPro; IPR007111; NACHT_NTPase.
DR Pfam; PF05729; NACHT; 1.
DR PROSITE; PS50837; NACHT; 1.
DR EMBL; AY338196; AAR03567.1; -.
DR InterPro; IPR007111; NACHT_NTPase.
DR Pfam; PF05729; NACHT; 1.
DR PROSITE; PS50837; NACHT; 1.
FT NON_TER 1
FT NON_TER 506
SQ SEQUENCE 506 AA; 58685 MW; BAA2A16BD583CAE4 CRC64;

Query Match 27.3%; Score 1493.5; DB 2; Length 506;
Best Local Similarity 55.3%; Pred. No. 8.4e-98;
Matches 281; Conservative 97; Mismatches 117; Indels 13; Gaps 4;

QY 132 RLLLVKHSNPMVOQQQLDTCGRHARTVGHQAQSPKIETLFFPDEERPEPPRTVMVMOGA 191
Db 1 RRLVKEHRSQOEREQELLAIK--TKTCEPSPVPIKWEMLFPDDHSEPHVTVVFOGA 58
QY 192 AGIGKSMIAHKVMDWADGKLFQGRFDYLFYINCREMNQSAATSCSMQDLIFSCWPEPSAP 251
Db 59 AGIGKTLARKIMLDWASGTLYQDRFDYLFYIHCREVS-LVTQSLGDLIVSCCPDPNPP 117
QY 252 LOELIRVPERLLFIIDGDELKPSFHDQPGWCLCWEKRPTELLINSIRKLLPELSL 311
Db 118 IHKIVRKPRIILFMDGDFDELQAFDEHIGPLCTDWQKAERGDIILSSIRKLLPEASL 177
QY 312 LITRPTALEKHLRLLHPRHVEILGFSEARKEYFYKYFHNAEQAGOVNRYVRDNEPLF 371
Db 178 LITRPALEKQLHLLDHPHVEILGFSEARKEYFFKYFSDQAARAAASLIQNEVLV 237
QY 372 TMCVPLVCWVCTCLOQLEGGLRQTSRTTAVTMYLLSLMQPKGAPRLQPPNQ 431
Db 238 TMCVPLVCWVCTGLQKQMESGKSLAQTSTTAVTVVFLSSLLQPRGGSQEHRLCAHL 297
QY 432 RGLCSLAADGLWNOKILFEEODLRKHGLDGEDVSFAFLNNIIFOKINCERYYSFIHLSFO 491
Db 298 WGLCSLAADGIWNOKILFEEODLRKHGLDGEDVSFAFLNNIIFOKINCERYYSFIHLSFO 357
QY 492 EFFAAMYIILDEGGAG-----PDQDVTLLTEYAFSERSFLALTSRFLGLNNE 542
Db 358 EFFAAMYILLEEKERRTNIPGSRLLKPSRDVTVLENYGKFEKGYLIFVVRFLGLVNO 417
QY 543 ETRSHLEKSLCWKSPHIMDLLOWIOSKAGSDGSTLQGSLEFFSCLYEIQBEETFOQA 602
Db 418 ERTSYLEKSLCKSKISQIRLELLKWIEVAKAKKLIQPSQLELFLCYEMQEEDFVQRA 477
QY 603 LSHFQVIVVSNIAKMEHVMVSSFCCLKRC 630
Db 478 MDYFPKIEI-NLSTRMDHVSVSSFCIENC 504

RESULT 10
Q68U59 PRELIMINARY; PRT; 506 AA.
ID Q68U59
AC Q68U59;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE Cryopyrin (Fragment).
GN Name=CIAS1;
OS Macaca sylvanus (Barbary ape).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheciniae; Macaca.
OX NCBI_TaxID=9546;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15302403; DOI=10.1016/j.gene.2004.05.002;
RA Anderson J.P., Mueller J.L., Rosengren S., Boyle D.L., Schaner P.,
RA Cannon S.B., Goodyear C.S., Hoffman H.M.;
RT "Structural, expression, and evolutionary analysis of mouse CIAS1.";
RL Gene 338:25-34(2004).
DR EMBL; AY338197; AAR03566.1; -.
DR EMBL; AY338197; AAR03566.1; -.

```

```

RT "Structural, expression, and evolutionary analysis of mouse CIAS1.";
RL Gene 338:25-34(2004).
DR EMBL; AY338198; AAR03567.1; -.
DR InterPro; IPR007111; NACHT_NTPase.
DR Pfam; PF05729; NACHT; 1.
DR PROSITE; PS50837; NACHT; 1.
FT NON_TER 1
FT NON_TER 506
SQ SEQUENCE 506 AA; 58812 MW; F8AC1357A5087C5E CRC64;

Query Match 27.0%; Score 1475.5; DB 2; Length 506;
Best Local Similarity 54.7%; Pred. No. 1.6e-96;
Matches 278; Conservative 98; Mismatches 119; Indels 13; Gaps 4;

QY 132 RLLLVKHSNPMVOQQQLDTCGRHARTVGHQAQSPKIETLFFPDEERPEPPRTVMVMOGA 191
Db 1 RRLVKEHRSQOEREQELLAIK--TKTCEPSPVPIKWEMLFPDDHSEPHVTVVFOGA 58
QY 192 AGIGKSMIAHKVMDWADGKLFQGRFDYLFYINCREMNQSAATSCSMQDLIFSCWPEPSAP 251
Db 59 AGIGKTLARKIMLDWASGTLYQDRFDYLFYIHCREVS-LVTQSLGDLIVSCCPDPNPP 117
QY 252 LOELIRVPERLLFIIDGDELKPSFHDQPGWCLCWEKRPTELLINSIRKLLPELSL 311
Db 118 IRKIVSKPSRIILFMDGDFDELQAFDEHIGPLCTDWQKAERGDIILSSIRKLLPEASL 177
QY 312 LITRPTALEKHLRLLHPRHVEILGFSEARKEYFYKYFHNAEQAGOVNRYVRDNEPLF 371
Db 178 LITRPALEKQLHLLDHPHVEILGFSEARKEYFFKYFSDQAARAAASLIQNEVLV 237
QY 372 TMCVPLVCWVCTCLOQLEGGLRQTSRTTAVTMYLLSLMQPKGAPRLQPPNQ 431
Db 238 TMCVPLVCWVCTGLQKQMESGKSLAQTSTTAVTVVFLSSLLQPRGGSQEHRLCAHL 297
QY 432 RGLCSLAADGLWNOKILFEEODLRKHGLDGEDVSFAFLNNIIFOKINCERYYSFIHLSFO 491
Db 298 WGLCSLAADGIWNOKILFEEODLRKHGLDGEDVSFAFLNNIIFOKINCERYYSFIHLSFO 357
QY 492 EFFAAMYIILDEGGAG-----PDQDVTLLTEYAFSERSFLALTSRFLGLNNE 542
Db 358 EFFAAMYILLEEKERRTNIPGSRLLKPSRDVTVLENYGKFEKGYLIFVVRFLGLVNO 417
QY 543 ETRSHLEKSLCWKSPHIMDLLOWIOSKAGSDGSTLQGSLEFFSCLYEIQBEETFOQA 602
Db 418 ERTSYLEKSLCKSKISQIRLELLKWIEVAKAKKLIQPSQLELFLCYEMQEEDFVQRA 477
QY 603 LSHFQVIVVSNIAKMEHVMVSSFCCLKRC 630
Db 478 MDYFPKIEI-NLSTRMDHVSVSSFCIENC 504

RESULT 11
Q68U60 PRELIMINARY; PRT; 506 AA.
ID Q68U60
AC Q68U60;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Cryopyrin (Fragment).
GN Name=CIAS1;
OS Theropithecus gelada (Gelada baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheciniae; Theropithecus.
OX NCBI_TaxID=9565;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15302403; DOI=10.1016/j.gene.2004.05.002;
RA Anderson J.P., Mueller J.L., Rosengren S., Boyle D.L., Schaner P.,
RA Cannon S.B., Goodyear C.S., Hoffman H.M.;
RT "Structural, expression, and evolutionary analysis of mouse CIAS1.";
RL Gene 338:25-34(2004).
DR EMBL; AY338197; AAR03566.1; -.
DR EMBL; AY338197; AAR03566.1; -.

```

InterPro; IPRO07111; NACHT_NTPase.
Pfam; PF05729; NACHT; 1.
DR PROSITE; PS50837; NACHT; 1.
FT NON_TER 1
SQ SEQUENCE 506 AA; 58737 MW; 3355EE2362C1D008 CRC64;
Query Match 27.0%; Score 1475.5; DB 2; Length 506;
Best Local Similarity 54.7%; Pred. No. 1.6e-96;
Matches 278; Conservative 98; Mismatches 119; Indels 13; Gaps 4;

QY 132 RLILVKEHNPQMVOQQOQLDTCGRGHARTVGHQAASPIKTIETLFEPEDEERPEPRVTVMOGA 191
||| : ||| :
Db 1 RLILVKEHNSQOEREHELLAIGK--TKTWESVPSPDKMELLFDPDEHSEPHVTVFQGA 58
||| : ||| :
QY 192 AGIGKMLAHKWMLDWADGKLFGQRFDFYLFIYNCREMNOSATECSMDLI FSCWPESAP 251
||| : ||| :
Db 59 AGIGKTILARKIMLDWASGTLYQDRFDYLFYIHCREVS-LVTRSLGDLINSCCPDNPP 117
||| : ||| :
QY 252 LQELIRVPERLLFIIDGDELKPSFHDPOGPWCCLWECKRPTTELLNLSLRKKLLPELSL 311
||| : ||| :
Db 118 IRKIVSKSRILLFLMDGDFDELQCAFDEHGIPCTDQKAERGDI LLSSLRKKLLPEASL 177
||| : ||| :
QY 312 LIITRTALEKHLRLLEHPRHVEIILGFSEAERKEYFYKFHNAQAGOVFNVRDNEPLF 371
||| : ||| :
Db 178 LIITRVPALEKLOHLLDHPRHVEIILGFSEAERKEYFYKFSDAQARAASFSLIQENSVLF 237
||| : ||| :
QY 372 TMCFFVLVCWVCTCLOQLEGGLLROTSTRITTA VMYLXLLSMQPKGAPRIQPPNQ 431
||| : ||| :
Db 238 TMCFFVLVCWVCTGLKQMESGKS LAQTSTTTAVTIFLSSLLQPRGGQEHLCAHL 297
||| : ||| :
QY 432 RGLCSLAADGLNWOKILFEEODLRKHGLDGEDVSAFLNNVIFOKDINCERYYSIIHLSFQ 491
||| : ||| :
Db 298 WGLCSLAADGIWNOKILFEESDLRNHGLQADVSAFLRMNLFOKEVDCEKYSIHMTFQ 357
||| : ||| :
QY 492 EFPAAMYIILDEGEAG-----PDQDVTRLLEYAFSERSFALTSRFLUGLLNE 542
||||| : ||| :
Db 358 EFPAAMYILLEBEKEGRTNVPGSLKLP SRDVTVLLENYGKFKGYLIFVVRFLGLVNQ 417
||| : ||| :
QY 543 ETRSHLEKSLCWKVSPIHKMDLLOWIOSKAQSGSTLQOGLSFPCLYEIOEBEFTQOA 602
||| : ||| :
Db 418 ERTCYLEKLSKISQIIRLELLKMWIEVKAKKLTQPSOLEIFYCLYEMQBEDVFQRA 477
||| : ||| :
QY 603 LSHFQVI VSNIA SKMEHVMSVFCCLKRC 630
: : | :
Db 478 MDYFPKIEI-NLSTRMDHVVSFCIENC 504

RESULT 12
Q68U56 PRELIMINARY; PRT; 500 AA.
ID Q68U56 AC Q68U56
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DE Cryopyrin (Fragment).
GN Names=CIAS1;
OS Alouatta belzebul (Black-and-red howler monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Alouattinae;
OC Alouatta.
ON NCBI_TaxId=30590;
RX PubMed=15302403; DOI=10.1016/j.gene.2004.05.002;
RA Anderson J.P., Mueller J.L., Rosengren S., Boyle D.L., Schaner P.,
Cannon S.B., Goodyear C.S., Hoffman H.M.;
RT "Structural, expression, and evolutionary analysis of mouse CIAS1.";
RL Gene 338:25-34(2004).
DR EMBL; AY338201; AAR03570.1; --
DR Pfam; PF05729; NACHT; 1.
DR InterPro; IPR007111; NACHT_NTPase.
FT NON_TER 1
SQ SEQUENCE 506 AA; 57708 MW; 055D921C39DDB6D7 CRC64;

Query Match 26.9%; Score 1471; DB 2; Length 500;
 Best Local Similarity 55.9%; Pred. No. 3.3e-96;
 Matches 281; Conservative 89; Mismatches 119; Indels 14; Gaps 4;

Qy 133 LLLVKEHSNPMQVQOQLDGTGRHARTVGHQASPIKIETLFPDDEBERPPPTVVMQGA 192
 Db 2 LRLIKEHSRQOEREHELLAIGK--TKTCSVPSPINMELLFPDDEHSEFVHTVVVFOGA 59

Qy 193 GICKSLAHKMLDWDAGKLFQGRFDYLFYINCREMNSQATSCSMODLIFSCWPBPSPAPL 252
 Db 60 GICKTILARKIMLDWASGTYLQDRFDYLFYIHCRELS-LVTQSLGDLIMSCCPDPNPPI 118

Qy 253 QELIRVPERLLFIIDGDELKPSFDPQGWCLCWEKEKPTTELLNSLIRKLLPELSL 312
 Db 119 HKIVRPSRILFLMDGDFDELQAFDEHIGPLCTNWKABERGDLSSLLIRKLLPEASL 178

Qy 313 ITRTPRALEKHLLEHPRHVEILGFSEARKEYFYKFFHNAEQAGQVFNVDNEPLFT 372
 Db 179 ITRTPVALEKLOHLLDHPHVEILGFSEARKEYFYKFFSDEAQPRAAFSLQENEVLF 238

Qy 373 MCFVPLVWVWVCTCLOOQLEGGGLLQRTSRTTAVTNMVLKLLSQPKGAPRIQPPNOR 432
 Db 239 MCFIPLVCWIVCTGLKQOMESGKSLAQTSTTTAVTFFLSSLLQPRGGSQEHPLSAHLW 298

Qy 433 GLCSLAADGIWNNOKILFEEODLRKHGDLGDSVAFINMNIPOKDCINERYYSFIHLSFOE 492
 Db 299 GLCSLAADGIWNNOKILFEEODLRKHGDLGDSVAFINMNIPOKDCINERYYSFIHMTFOE 358

Qy 493 FFAAMYVILDEGEGAG-----PDQDVTRLLTEYAFSERSFALTSRFLFGLNE 542
 Db 359 FFAAMYVILDEGEGAG-----PDQDVTRLLTEYAFSERSFALTSRFLFGLVNO 418

Qy 543 ETRSHLEKSLCWKVSPIKMDLLOMIQSKAQSDGSTLQOQSLFFSCLYEIQEEFIQQA 602
 Db 419 ERTSYLEKSLCKSVSQIRLELLKWIYKAKKAKKQIQPSQLELYFYCYEMQEAFVORA 478

Qy 603 LSHFQVIVVSNIAKMEHWSSP 625
 Db 479 MDYFPRIE-ILSTRMDHVSSP 500

RESULT 14
 Q68U57 PRELIMINARY; PRT; 506 AA.
 ID AC Q68U57
 DT 25-OCT-2004 (TREMELrel. 28, Created)
 DT 25-OCT-2004 (TREMELrel. 28, Last sequence update)
 DE Cycopryrin (Fragment).
 GN Name=CIA1;
 OS Colobus guereza (Black-and-white colobus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Colobinae;
 OC Colobus.
 OX NCBI_TaxID=33548;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15302403; DOI=10.1016/j.gene.2004.05.002;
 RA Anderson J.P., Mueller J.L., Rosengren S., Boyle D.L., Schaner P.,
 RA Cannon S.B., Goodyear C.S., Hoffman H.M.;
 RT "Structural, expression, and evolutionary analysis of mouse CIA1.";
 RL Gene 338:25-34(2004).
 DR EMBL; AY338200; AAR03569.1; --
 DR InterPro; IPR007111; NACHT_NTPase.
 DR Pfam; PF05729; NACHT; 1.
 DR PROSITE; PS50837; NACHT; 1.
 FT NON_TER 1
 FT NON_TER 506
 SQ SEQUENCE 506 AA; 58700 MW; 435425BC125BEAD4 CRC64;

Query Match 26.9%; Score 1470.5; DB 2; Length 506;
 Best Local Similarity 54.5%; Pred. No. 3.7e-96;
 Matches 277; Conservative 98; Mismatches 120; Indels 13; Gaps 4;

Qy 132 RLLVKEHSNPMQVQOQLDGTGRHARTVGHQASPIKIETLFPDDEBERPPPTVVMQGA 191
 Db 1 RRLIKEHSRQOEREHELLAIGK--TKTCSVPSPINMELLFPDDEHSEFVHTVVVFOGA 58

Qy 192 AGICKSLAHKMLDWDAGKLFQGRFDYLFYINCREMNSQATSCSMODLIFSCWPBPSPAP 251
 Db 59 AGICKTILARKIMLDWASGTYLQDRFDYLFYIHCREVS-LVTQSLGDLIMSCCPDPNP 117

Qy 252 QELIRVPERLLFIIDGDELKPSFDPQGWCLCWEKEKPTTELLNSLIRKLLPELSL 311
 Db 118 IRKIVSPSRILFLMDGDFDELQAFDEHIGPLCTNWKABERGDLSSLLIRKLLPEASL 177

Qy 312 ITRTPRALEKHLLEHPRHVEILGFSEARKEYFYKFFHNAEQAGQVFNVDNEPLF 371
 Db 178 ITRTPVALEKLOHLLDHPHVEILGFSEARKEYFYKFFSDEAQARAAFSLIQENEVLF 237

Qy 372 TMCVPLVWVWVCTCLOOQLEGGGLLQRTSRTTAVTNMVLKLLSQPKGAPRIQPPNQ 431
 Db 238 TMCVPLVWVWVCTCLOOQLEGGGLLQRTSRTTAVTNMVLKLLSQPKGAPRIQPPNQ 297

Qy 432 RGLCSLAADGIWNNOKILFEEODLRKHGDLGDSVAFINMNIPOKDCINERYYSFIHLSFO 491
 Db 298 WGLCSLAADGIWNNOKILFEEODLRKHGDLGDSVAFINMNIPOKDCINERYYSFIHMTFO 357

Qy 492 EFAAMYVILDEGEGAG-----PDQDVTRLLTEYAFSERSFALTSRFLFGLNE 542
 Db 358 EFAAMYVILDEGEGAG-----PDQDVTRLLTEYAFSERSFALTSRFLFGLVNO 417

Qy 543 ETRSHLEKSLCWKVSPIKMDLLOMIQSKAQSDGSTLQOQSLFFSCLYEIQEEFIQQA 602
 Db 418 ERTCYLEKSLCKSVSQIRLELLKWIYKAKKAKKQIQPSQLELYFYCYEMQEEDFVORA 477

Qy 603 LSHFQVIVVSNIAKMEHWSSP 630
 Db 478 MDYFPKIEI-ILSTRMDHVSSP 504

RESULT 15
 Q68U58 PRELIMINARY; PRT; 506 AA.
 ID AC Q68U58
 DT 25-OCT-2004 (TREMELrel. 28, Created)
 DT 25-OCT-2004 (TREMELrel. 28, Last sequence update)
 DE Cycopryrin (Fragment).
 GN Name=CIA1;
 OS Cercopithecius cephus (moustached monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Cercopithecius.
 OX NCBI_TaxID=9535;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15302403; DOI=10.1016/j.gene.2004.05.002;
 RA Anderson J.P., Mueller J.L., Rosengren S., Boyle D.L., Schaner P.,
 RA Cannon S.B., Goodyear C.S., Hoffman H.M.;
 RT "Structural, expression, and evolutionary analysis of mouse CIA1.";
 RL Gene 338:25-34(2004).
 DR EMBL; AY338199; AAR03568.1; --
 DR InterPro; IPR007111; NACHT_NTPase.
 DR Pfam; PF05729; NACHT; 1.
 DR PROSITE; PS50837; NACHT; 1.
 FT NON_TER 1
 FT NON_TER 506
 SQ SEQUENCE 506 AA; 58807 MW; 3BD8387BC8ACBB63 CRC64;

Query Match 26.8%; Score 1468.5; DB 2; Length 506;
 Best Local Similarity 54.5%; Pred. No. 5.1e-96;
 Matches 277; Conservative 98; Mismatches 120; Indels 13; Gaps 4;

Qy 132 RLLVKEHSNPMQVQOQLDGTGRHARTVGHQASPIKIETLFPDDEBERPPPTVVMQGA 191
 Db 1 RRLIKEHSRQOEREHELLAIGK--TKTCSVPSPINMELLFPDDEHSEFVHTVVVFOGA 58

This Page Blank (uspto)

| Result No. | Score | Query | | DB | ID | Description |
|------------|-------|-------|--------|----|--------------------|-------------------|
| | | Match | Length | | | |
| 1 | 1375 | 25.1 | 1454 | 4 | US-09-388-221B-10 | Sequence 10, Appl |
| 2 | 1375 | 25.1 | 1473 | 4 | US-09-388-221B-2 | Sequence 2, Appl |
| 3 | 1360 | 24.9 | 1399 | 4 | US-09-388-221B-4 | Sequence 4, Appl |
| 4 | 1360 | 24.9 | 1424 | 4 | US-09-388-221B-12 | Sequence 12, Appl |
| 5 | 1360 | 24.9 | 1443 | 4 | US-09-388-221B-6 | Sequence 6, Appl |
| 6 | 723 | 13.2 | 461 | 2 | US-08-910-731-6 | Sequence 6, Appl |
| 7 | 717 | 13.1 | 456 | 2 | US-08-910-731-4 | Sequence 4, Appl |
| 8 | 717 | 13.1 | 456 | 2 | US-08-795-395-4 | Sequence 4, Appl |
| 9 | 712 | 13.0 | 456 | 2 | US-08-910-731-8 | Sequence 8, Appl |
| 10 | 709 | 13.0 | 456 | 2 | US-08-910-731-2 | Sequence 2, Appl |
| 11 | 709 | 13.0 | 456 | 2 | US-08-795-395-2 | Sequence 2, Appl |
| 12 | 529 | 9.7 | 1013 | 4 | US-10-014-269-3 | Sequence 3, Appl |
| 13 | 529 | 9.7 | 1040 | 4 | US-10-014-269-2 | Sequence 2, Appl |
| 14 | 529 | 9.7 | 1043 | 4 | US-09-949-016-9080 | Sequence 9080, Ap |
| 15 | 526 | 9.6 | 1040 | 4 | US-10-014-269-4 | Sequence 4, Appl |
| 16 | 522 | 9.5 | 953 | 3 | US-09-099-041A-8 | Sequence 8, Appl |
| 17 | 522 | 9.5 | 953 | 3 | US-09-245-281-8 | Sequence 8, Appl |
| 18 | 522 | 9.5 | 953 | 4 | US-09-207-359B-8 | Sequence 8, Appl |
| 19 | 522 | 9.5 | 953 | 4 | US-09-340-620A-8 | Sequence 8, Appl |
| 20 | 522 | 9.5 | 953 | 4 | US-09-865-364-8 | Sequence 8, Appl |
| 21 | 511 | 9.1 | 1007 | 4 | US-10-014-269-34 | Sequence 34, Appl |
| 22 | 498.5 | 9.1 | 209 | 4 | US-09-388-221B-19 | Sequence 19, Appl |
| 23 | 464.5 | 8.5 | 953 | 3 | US-09-245-281-43 | Sequence 43, Appl |
| 24 | 464.5 | 8.5 | 953 | 4 | US-09-207-359B-43 | Sequence 43, Appl |
| 25 | 464.5 | 8.5 | 953 | 4 | US-09-340-620A-43 | Sequence 43, Appl |
| 26 | 464.5 | 8.5 | 953 | 4 | US-09-865-364-43 | Sequence 43, Appl |
| 27 | 455.5 | 8.3 | 966 | 4 | US-09-207-359B-47 | Sequence 47, Appl |

Db 363 HVFYFSCRELAQSKV-VSLAELIGKDGATAPAPIRQILSRPERLLFILGDVDEPGWVLQE 421
Qy 279 PQGFWCLCWEKRPTELLNSLRKXLLPELSLLITRPTALEKLRHLEHPRHVEILGF 338
Db 422 PSSELCHWSQOPADALLGSLGKTLPEASPLITARTALQNLIPSEQARWVEVLGF 481
Qy 339 SEAEKKEYFYKPHNAEQOVNRYVDRNEPLFTMCFVPLVCWVCTCLQQOEGGGLLR 398
Db 482 SSSRKEYFYRFTDERQAIRAFRLVKSNEKELWALCLVPWVSLACTCLMQQKKEKLT 541
Qy 399 QTSRTTAVMYLLSMLQPKGAPRLQPPNQRGLCSLAADGLMNQKILFEEQDLRKHG 458
Db 542 LTKTITTLCHYLAQALQALPGQL-----RDLCSLAEEGIWQKTLFSPDDLRKHG 595
Qy 459 LDGEDVSALNMNIFQKINCERYYSFIHLSFOEFPAAMYIILDEGEGAGPDQ----DV 514
Db 596 LDGAIISTFLKMGILQEH-PIPLSYSFHLCFOEFPAAMSYVL-EDEKRGKHSNCIIDL 653
Qy 515 TRLLTEYAFSERSFLALTSRFLGLNNEETRSHEKSLCWKSPHIMKMDLLOWIQSKAQS 574
Db 654 EKTLEYAGI-HGLFGASTTRFLGLLSDEGEREMENIFHCRLSQ--GRNLMQWVPSLQ-- 708
Qy 575 DGSSTLQGSLEFFSCLYEIOEEFIIQOALSHFQVIVVSNIAKQVHMHVSSFCIKRCRSQA 634
Db 709 --LLQPHSLSLHCLYETRNKTLFTQVMAHFEEMGMC-VETDMELLVCTFCIKFSRHVK 765
Qy 635 VILHYGATYSADGEDRARCAGATLLVQLRPERTVLLDAYSEHLAAALCTNPNLIELSL 694
Db 766 KLQLI-----EGRQHRSTWSPWVLF--WVPVTDAYWQLFSLVKVTRNLKELDL 815
Qy 695 YRNALSGRVKLLCOGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSGN 754
Db 816 SGNLSHSAVKSCKTLRPRCLLETRLAGCGLTAECDCKDLAFGLRANQTLIELDLSFN 875
Qy 755 GVGFPGMMLCEGLRHPQCRLOMTQLRKQLESQACQEMASVLGTNPHILVELDLTGNALE 814
Db 876 VLTDAAGAKHLQRLQPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKELDLQNNLD 935
Qy 815 DLGRLILCOGLRHPVCRRLTLMWKICRLTAACDELASTLSVNSQSLRELDLSNEL---- 870
Db 936 DVGVRLLCEGLRHPACKLIRLGL-----DQTLSDMRQELRALQEQKPOLLIIFS 985
Qy 871 -GDLGVLLLCBGL-----RHPTCKLQTLRLGICRLGSAACEGLSVVLQANHNLRELDLS 923
Db 986 RRKPSVMTPEGLDTGEMSNSTSSLRQRLGSRASASHV-----AQANLKLLDVS 1035

RESULT 2

US-09-388-221B-2
; Sequence 2, Application US/09388221B
; Patent No. 6818750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6818750el Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1473
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221B-2
Query Match 25.1%; Score 1375; DB 4; Length 1473;
Best Local Similarity 33.9%; Pred. No. 9.4e-126;
Matches 366; Conservative 138; Mismatches 357; Indels 218; Gaps 23;
Qy 12 RLSTVLEEAEVLEKFKLYLCTA--TELGEKIPWGSMEKAGPLEMAQLIITHFGPEEA 69
Db 8 RLACYLEFLKKELEKFEQILLANKAHSRSSGETP-AQPEKTSNGNEVASYLVAQYGEORA 66

Qy 70 WRLALSTPERINRKDLWERGQ-----EDLVRDPQTYRDYV----- 106
Db 67 WDLALHTWEQWGLSRSLCAQACQAGHAGSPSPYSPSEPHLGSPSQPTSTAVLMPWIHELPA 126
Qy 107 -----RRKPRIMEDRNARLGEVCUNLSHRTRYTRLLLVKEHSNPMQVQ----- 146
Db 127 GCTQGSERRVLQPLDTSGRWRREISASLLYQALFSPDHESPSQESNAPTSTAVLGSW 186
Qy 147 ----- 146
Db 187 GSPPOPSLAPREQAPGTQWFLDETSGTYIYTEREREREKSEKGRPPWAAVVGTPPQAHT 246
Qy 147 -----QQLDTRGRHARTV----- 160
Db 247 SLQPHHHHPWPSVRESLCSTWPKNEDFNQKFTQLLLIQRPHPRSQDPLVKRSPDYVEE 306
Qy 161 --GHQASIKTETLEPDEERPEPRVTVMOGAGIGKSLAHKMLMDADGKLFQGRFD 218
Db 307 NRGHJL---TEIRDLPGLDQOE-PRVILQOAGIGKSTLARQVKEAMGRGQLYGDRAFO 362
Qy 219 YLFYINCREMNOQATECSMDLIIFSCWPEPSAPIQELIRVPERLLFIIDGDELKPSPHD 278
Db 363 HVFYFSCRELAQSKV-VSLAELIGKDGATAPAPIRQILSRPERLLFILGDVDEPGWVLQE 421
Qy 279 PQGFWCLCWEKRPTELLNSLRKXLLPELSLLITRPTALEKLRHLEHPRHVEILGF 338
Db 422 PSSELCHWSQOPADALLGSLGKTLPEASPLITARTALQNLIPSEQARWVEVLGF 481
Qy 339 SEAEKKEYFYKPHNAEQOVNRYVDRNEPLFTMCFVPLVCWVCTCLQQOEGGGLLR 398
Db 482 SSSRKEYFYRFTDERQAIRAFRLVKSNEKELWALCLVPWVSLACTCLMQQKKEKLT 541
Qy 399 QTSRTTAVMYLLSMLQPKGAPRLQPPNQRGLCSLAADGLMNQKILFEEQDLRKHG 458
Db 542 LTKTITTLCHYLAQALQALPGQL-----RDLCSLAEEGIWQKTLFSPDDLRKHG 595
Qy 459 LDGEDVSALNMNIFQKINCERYYSFIHLSFOEFPAAMYIILDEGEGAGPDQ----DV 514
Db 596 LDGAIISTFLKMGILQEH-PIPLSYSFHLCFOEFPAAMSYVL-EDEKRGKHSNCIIDL 653
Qy 515 TRLLTEYAFSERSFLALTSRFLGLNNEETRSHEKSLCWKSPHIMKMDLLOWIQSKAQS 574
Db 654 EKTLEYAGI-HGLFGASTTRFLGLLSDEGEREMENIFHCRLSQ--GRNLMQWVPSLQ-- 708
Qy 575 DGSSTLQGSLEFFSCLYEIOEEFIIQOALSHFQVIVVSNIAKQVHMHVSSFCIKRCRSQA 634
Db 709 --LLQPHSLSLHCLYETRNKTLFTQVMAHFEEMGMC-VETDMELLVCTFCIKFSRHVK 765
Qy 635 VILHYGATYSADGEDRARCAGATLLVQLRPERTVLLDAYSEHLAAALCTNPNLIELSL 694
Db 766 KLQLI-----EGRQHRSTWSPWVLF--WVPVTDAYWQLFSLVKVTRNLKELDL 815
Qy 695 YRNALSGRVKLLCOGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSGN 754
Db 816 SGNLSHSAVKSCKTLRPRCLLETRLAGCGLTAECDCKDLAFGLRANQTLIELDLSFN 875
Qy 755 GVGFPGMMLCEGLRHPQCRLOMTQLRKQLESQACQEMASVLGTNPHILVELDLTGNALE 814
Db 876 VLTDAAGAKHLQRLQPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKELDLQNNLD 935
Qy 815 DLGRLILCOGLRHPVCRRLTLMWKICRLTAACDELASTLSVNSQSLRELDLSNEL---- 870
Db 936 DVGVRLLCEGLRHPACKLIRLGL-----DQTLSDMRQELRALQEQKPOLLIIFS 985
Qy 871 -GDLGVLLLCBGL-----RHPTCKLQTLRLGICRLGSAACEGLSVVLQANHNLRELDLS 923
Db 986 RRKPSVMTPEGLDTGEMSNSTSSLRQRLGSRASASHV-----AQANLKLLDVS 1035

RESULT 3

US-09-388-221B-4
; Sequence 4, Application US/09388221B
; Patent No. 6818750

| | | | |
|---|-----|--|------|
| QY | 339 | SEAEKKEYFYKYFHNAEQAGOVNRYVDRNEPLFTMCVFPLVCWVCTCLOQLEGGGLR | 398 |
| Db | 482 | SESRKEYFYRYFTDERQAIRAFRLVKSNEKELWALCLVPWVSWLACTCLLMQOMKREKLT | 541 |
| QY | 399 | QTSRTTAVVYMLYLLSMQPKGAPRLQPPNORGLCSLAADGLWNOKILFEEQDLRKHG | 458 |
| Db | 542 | LTSKTTTTLCHYLAAQALQAPLGPOL-----RDLCSLAAGIWKQKTLFSPDDLKHG | 595 |
| QY | 459 | LDGEDVSAPLNNMIFOKDINCERYYSFIHLSPOEFFAAMYIILDEGEAGDPQ---- | 514 |
| Db | 596 | LDGAIISHTFUMGLQLQH-PIPLSYSFHLCFOEFFAAMSYVL-EDEKGRGKHSNCIIDL | 653 |
| QY | 515 | TRLITEYAFSERSFALTSRFLGLNNEETRSLSKSLCWKVSPPHIMDLOWIQSKAOS | 574 |
| Db | 654 | EKTLVAYGI-HGLFGASTTRFLGLLSDEGEREMENIFHCRLSQ--GRNLQWVPSLQ-- | 708 |
| QY | 575 | DGSTLOQGSLEFFSCLYIEEERFIOQALSHFQVIVVSNIAKWEHVMVSSFCIKRCSAQ | 634 |
| Db | 709 | --LLOPHSLESLSHCLYETRNKTLFTQVMAHFEEMGMC-VETDMELLVCTFCIKFSRHVK | 765 |
| QY | 635 | VLHYGATYSADGEDRARCAGAHTLVQLRPRTVLLDAYSEHLAAALCTNPNLIELSL | 694 |
| Db | 766 | KLQILI-----EGRQHRSTWSPVTVLFR--WVPVTDAYWQILFSLVKVTRNLKELD | 815 |
| QY | 695 | YRNALSGRGVKLLCOGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSGN | 754 |
| Db | 816 | SGNSLSHSAVKSCKTLRRPRCLLETFLRAGCGLTAEDCKDLAFGLRANQTLTELDLSFN | 875 |
| QY | 755 | GVGPPGMMLCEGLRHPQCRLOMIQLRKQLESAGCOEMASVLGTPHPLVELDLTGNALE | 814 |
| Db | 876 | VLTDAGAKHLQRLRQPSCKLQRLQVSCGLTSDCCODLASVLSASPSLKELDLQNNL | 935 |
| QY | 815 | DLGLRLCOGLRHPVCRILTLMWKICRLTAACDELASTLSVNSQSLRELDLSINELGDLG | 874 |
| Db | 936 | DVGVELLCEGLRHPACK-----LIRLGKPS | 960 |
| QY | 875 | VLLICEGL-----RHPTCKLQTLRLGICRLGSAACEGLSVVLQANHNLRELDLS | 923 |
| Db | 961 | VMTPTTEGLDTGEMSNSTSSLRQRLGSEAAASHV-----AQANLKLLDVS | 1005 |
| RESULT 5 | | | |
| US-09-388-221B-6 | | | |
| ; Sequence 6, Application US/09388221B | | | |
| ; Patent No. 6818750 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Reed, John C. | | | |
| ; TITLE OF INVENTION: No. 6818750el Card Proteins Involved in Cell Death Regulation | | | |
| ; FILE REFERENCE: P-LJ 3650 | | | |
| ; CURRENT APPLICATION NUMBER: US/09/388,221B | | | |
| ; NUMBER OF SEQ ID NOS: 30 | | | |
| ; SOFTWARE: Patent In Ver. 2.0 | | | |
| ; SEQ ID NO 6 | | | |
| ; LENGTH: 1443 | | | |
| ; TYPE: PRT | | | |
| ; ORGANISM: Homo sapiens | | | |
| US-09-388-221B-6 | | | |
| Query Match 24.9%; Score 1360; DB 4; Length 1443; | | | |
| Best Local Similarity 33.5%; Pred. No. 2.7e-124; | | | |
| Matches 360; Conservative 133; Mismatches 343; Indels 238; Gaps 22; | | | |
| QY | 12 | RLSTYLEEAEVLEKFKLYLGTAA--TELGEKIPWGSMEKAGPLEMAQLLTHFGPEEA | 69 |
| Db | 8 | RLACYLEFLKKEELKEFQLLANKAHSRSSGETP-AQPEKTSGEVASYLVAQYGEQRA | 66 |
| QY | 70 | WRALSTFERINKDLWERGQ-----EDLVDPQETRYDYV----- | 106 |
| Db | 67 | WDLALHTWEQWGLRSLCAQAGAGHSPSPFPSPHGLSPSQPTSTAVLPMWIHELPA | 126 |
| QY | 107 | -----RRKFLMEDNARLGCVNLSHRYFTLLLVKHSNPMQVQ----- | 146 |
| Db | 127 | GCTGGSERVLRQLPDTSGRRWREISASLLYQALPSPDHPESQSPNAPTSTAVLGSW | 186 |
| QY | 147 | ----- | 146 |
| Db | 187 | GSPQPSLAPREOAPCTOWPLDETSGIYYTEIRERERESEKGRPPWAAVGTTPQAHT | 246 |
| QY | 147 | -----OQLDTRGHARTV----- | 160 |
| Db | 247 | SLQPHHPWPSVRESLCTWPKNEDFNQFTQLLLQRPHPRSQDPLVKRSPWDYVEE | 306 |
| QY | 161 | --GHAQSPIKIEITLFEPEDEERPEPPRVVMQAGAGICKMLAKHVMLDWADGKLPGREFD | 218 |
| Db | 307 | NRGHL---IEIRDLFGGLDQE--PRIVILQGAAGICKSTARQVKEAWRGOLYGRFQ | 362 |
| QY | 219 | YLFVINCENMOSATESMODLIFSCWPESAPLOELIRVPERLLFTIDGFDLKPSPHD | 278 |
| Db | 363 | HVFYPSCHLAQSKV-VSLAELIKDGTATPAPRIOLSRPERLLFLDGVDPGWLQE | 421 |
| QY | 279 | POGWCLEWEKRPTELLNLSLRKLLPELSLITTRPTALEKHLRLEHPRHVEILGF | 338 |
| Db | 422 | PSSELCLHWSQPADALLGSLQKLTILPEASFLITARTTALQNLIPSEQARWVEVLG | 481 |
| QY | 339 | SEAEKKEYFYKYFHNAEQAGOVNRYVDRNEPLFTMCVFPLVCWVCTCLOQLEGGGLR | 398 |
| Db | 482 | SESRKEYFYRYFTDERQAIRAFRLVKSNEKELWALCLVPWVSWLACTCLLMQOMKREKLT | 541 |
| QY | 399 | QTSRTTAVVYMLYLLSMQPKGAPRLQPPNORGLCSLAADGLWNOKILFEEQDLRKHG | 458 |
| Db | 542 | LTSKTTTTLCHYLAAQALQAPLGPOL-----RDLCSLAAGIWKQKTLFSPDDLKHG | 595 |
| QY | 459 | LDGEDVSAPLNNMIFOKDINCERYYSFIHLSPOEFFAAMYIILDEGEAGDPQ---- | 514 |
| Db | 596 | LDGAIISHTFUMGLQLQH-PIPLSYSFHLCFOEFFAAMSYVL-EDEKGRGKHSNCIIDL | 653 |
| QY | 515 | TRLITEYAFSERSFALTSRFLGLNNEETRSLSKSLCWKVSPPHIMDLOWIQSKAOS | 574 |
| Db | 654 | EKTLVAYGI-HGLFGASTTRFLGLLSDEGEREMENIFHCRLSQ--GRNLQWVPSLQ-- | 708 |
| QY | 575 | DGSTLOQGSLEFFSCLYIEEERFIOQALSHFQVIVVSNIAKWEHVMVSSFCIKRCSAQ | 634 |
| Db | 709 | --LLOPHSLESLSHCLYETRNKTLFTQVMAHFEEMGMC-VETDMELLVCTFCIKFSRHVK | 765 |
| QY | 635 | VLHYGATYSADGEDRARCAGAHTLVQLRPRTVLLDAYSEHLAAALCTNPNLIELSL | 694 |
| Db | 766 | KLQILI-----EGRQHRSTWSPVTVLFR--WVPVTDAYWQILFSLVKVTRNLKELD | 815 |
| QY | 695 | YRNALSGRGVKLLCOGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSGN | 754 |
| Db | 816 | SGNSLSHSAVKSCKTLRRPRCLLETFLRAGCGLTAEDCKDLAFGLRANQTLTELDLSFN | 875 |
| QY | 755 | GVGPPGMMLCEGLRHPQCRLOMIQLRKQLESAGCOEMASVLGTPHPLVELDLTGNALE | 814 |
| Db | 876 | VLTDAGAKHLQRLRQPSCKLQRLQVSCGLTSDCCODLASVLSASPSLKELDLQNNL | 935 |
| QY | 815 | DLGLRLCOGLRHPVCRILTLMWKICRLTAACDELASTLSVNSQSLRELDLSINELGDLG | 874 |
| Db | 936 | DVGVELLCEGLRHPACK-----LIRLGKPS | 960 |
| QY | 875 | VLLICEGL-----RHPTCKLQTLRLGICRLGSAACEGLSVVLQANHNLRELDLS | 923 |
| Db | 961 | VMTPTTEGLDTGEMSNSTSSLRQRLGSEAAASHV-----AQANLKLLDVS | 1005 |
| RESULT 6 | | | |
| US-08-910-731-6 | | | |
| ; Sequence 6, Application US/08910731 | | | |
| ; Patent No. 5932440 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: CHATTERJEE, DEB K. | | | |
| ; APPLICANT: SHANDILYA, HARINI | | | |
| ; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof | | | |
| ; NUMBER OF SEQUENCES: 16 | | | |
| ; Sequence 6, Application US/08910731 | | | |
| ; Patent No. 5932440 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: CHATTERJEE, DEB K. | | | |
| ; APPLICANT: SHANDILYA, HARINI | | | |
| ; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof | | | |
| ; NUMBER OF SEQUENCES: 16 | | | |

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
;; CITY: WASHINGTON
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005-3934
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/910,731
;; FILING DATE: (Herewith)
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/795,395
;; FILING DATE: 04-FEB-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/794,546
;; FILING DATE: 03-FEB-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/024,057
;; FILING DATE: 16-AUG-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: ESMOND, ROBERT W.
;; REGISTRATION NUMBER: 32,893
;; REFERENCE/DOCKET NUMBER: 0942.3440003
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-371-2600
;; TELEFAX: 202-371-2540
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 461 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
US-08-910-731-6

Query Match 13.2%; Score 723; DB 2; Length 461;
Best Local Similarity 42.5%; Pred. No. 3.2e-62;
Matches 150; Conservative 64; Mismatches 139; Indels 0; Gaps 0;
Qy 671 LLDAYSEHLAAALCTNPNIJELSLYRNALSGRGVKKLCOGLRHPNCKLQNLRLKRCRISS 730
Db 40 LTEARCKDISSALRVNPAALNLSNELGVDGVHCVLQGLQTSCKIQKLSLQNCCLTG 99
Qy 731 SACEDLSAALIANKNLTRMDLSNGVGPFGMMLLCEGLRHPQCRQLQMIQLRKCOLESAC 790
Db 100 AGCGVLSSTLRTLPTLOELHLSNLDGAGLQALLCEGLLDPCQRLKLEYCYSLSAASC 159
Qy 791 QEMASVLGTNPVHVELDLTGNALEDGLRLCOGLRHPVCELRITLWIKICBLTAACDEL 850
Db 160 EPLASVLRAPDFKELTVSNNDINEAGRVLCQGLKDSPCQLEALKLESQVTSNCRDL 219
Qy 851 ASTLSVNQSLRELDLSNELGDLGVLLCEGLRHPTCKIQLTRIGICRLGSAACEGLSVV 910
Db 220 CGIVASVKSRLRELALGSKNGDGMALCPGLLPPSRRLRTLWECGITAAGCGDLCRV 279
Qy 911 LQANHNLRELDLSFNDLGDWGLWLLAEGLQHPACRLQKWLDSCLGTAKACENLYFTLGI 970
Db 280 LRAKESLKELSLGNELGDEGARLICETLLEPGQLESWLKVSFSAACCPHFSSVLAQ 339
Qy 971 NQTLTDLVLTNNALGDTGVRLCKRLSHPGCKLRLVLMFGMDLNKMTSHRLAA 1023
Db 340 NRFLLEQLISNNRLEDAGVRELCOGLQPGSVLRVLMADCDVSDSSCSLAA 392

RESULT 7
US-08-910-731-4
; Sequence 4, Application US/08910731

;; Patent No. 5932440
;; GENERAL INFORMATION:
;; APPLICANT: CHATTERJEE, DEB K.
;; APPLICANT: SHANDILYA, HARINI
;; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
;; NUMBER OF SEQUENCES: 16
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
;; CITY: WASHINGTON
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005-3934
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/910,731
;; FILING DATE: (Herewith)
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/795,395
;; FILING DATE: 04-FEB-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/794,546
;; FILING DATE: 03-FEB-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/024,057
;; FILING DATE: 16-AUG-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: ESMOND, ROBERT W.
;; REGISTRATION NUMBER: 32,893
;; REFERENCE/DOCKET NUMBER: 0942.3440003
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-371-2600
;; TELEFAX: 202-371-2540
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 456 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-910-731-4

Query Match 13.1%; Score 717; DB 2; Length 456;
Best Local Similarity 41.5%; Pred. No. 1.2e-61;
Matches 147; Conservative 66; Mismatches 141; Indels 0; Gaps 0;
Qy 671 LLDAYSEHLAAALCTNPNIJELSLYRNALSGRGVKKLCOGLRHPNCKLQNLRLKRCRISS 730
Db 35 LTVRCCKDISSALRVNPAALNLSNELGVDGVHCVLQGLQTSCKIQKLSLQNCCLTG 94
Qy 731 SACEDLSAALIANKNLTRMDLSNGVGPFGMMLLCEGLRHPQCRQLQMIQLRKCOLESAC 790
Db 95 AGCGVLPDVLRSLSSTLRELHNDNPLDGEGLKLLCEGLRDPQCRLEKLEYCYNLTATSC 154
Qy 791 QEMASVLGTNPVHVELDLTGNALEDGLRLCOGLRHPVCELRITLWIKICBLTAACDEL 850
Db 155 EPLASVLRVPDFKELTVSNNDINEAGIHTLCOGLKDSACQLESKLENCGITSANCKDL 214
Qy 851 ASTLSVNQSLRELDLSNELGDLGVLLCEGLRHPTCKIQLTRIGICRLGSAACEGLSVV 910
Db 215 CDVVASVKSRLRELALGSKNGDGMALCPGLLPPSRRLRTLWECGITAAGCGDLCRV 274
Qy 911 LQANHNLRELDLSFNDLGDWGLWLLAEGLQHPACRLQKWLDSCLGTAKACENLYFTLGI 970
Db 275 LRAKESLKELSLGNELKDEGAQLCELSLLEPGQLESWLKVSFSAACCPHFCSVLTK 334
Qy 971 NQTLTDLVLTNNALGDTGVRLCKRLSHPGCKLRLVLMFGMDLNKMTSHRLAA 1024
Db 335 NSSLFEQLQSNPLGDSGVVELCKALGYPTVLRVLMADCDVSDSSCSLAA 388

RESULT 8

US-08-795-395-4
; Sequence 4, Application US/08795395
; Patent No. 5965399
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Cloning and Expression of Rat Liver and
; TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,395
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-795-395-4

Query Match 13.1%; Score 717; DB 2; Length 456;
Best Local Similarity 41.5%; Pred. No. 1.2e-61;
Matches 147; Conservative 66; Mismatches 141; Indels 0; Gaps 0;

| | | | | | | |
|----|-----|---------------------|-------------------|--------------------|-----------------------|-----------|
| Qy | 671 | LLDAYSEHLAAALCTNPNI | ELSLYRNALSGRGVLL | COGLRHPNCKQLNLRKCR | ISS 730 | |
| Db | 35 | LTEVRCKDIRSAIQANPA | LTSELRTELGDAGVGL | VLQGLQNTCKIQKLSLQ | NCSLTE 94 | |
| Qy | 731 | SACEDLSAALIANKNLT | RMDSLGNVGFPGNMILL | CEGLRHPNCKQLNLRKCR | ISS 790 | |
| Db | 95 | AGCGVLPDVRSLSTLR | ELHNDPLNDPLGDEGL | KLLCEGLRDPQCRLEK | LEYCNLTATSC 154 | |
| Qy | 791 | QEMASVLGTNPVHVELD | ITGNALDELGLRL | COGLRHPNCKQLNLRKCR | ISS 850 | |
| Db | 155 | EPLASVLVRKDPFKE | LVLSNNDPHEAGITHT | LCQGLKDSACQLES | SLKLENCGITSANCKDL 214 | |
| Qy | 851 | ASTLSVNSQSLRELD | LSNELDGLVLL | CEGLRHPNCKQLNLRKCR | ISS 910 | |
| Db | 215 | CDVASKASLQELD | LSGNKNGTGAALCS | GILLPSCRLTILWDCD | VTABGCKDLCRV 274 | |
| Qy | 911 | LQAHNLRDELDSFND | LGDWGLWLAEGLO | HPACRLQKWLDS | CGLTAKACENLYFTLGI 970 | |
| Db | 275 | LRAQSKLSL | AGNELKDEGAQL | CELSLEPGCQLES | LVKTCSTLTAASCPHFC | SVLTK 334 |
| Qy | 971 | NOTLTDLYLTNAL | GDTGVRLICKEL | SHPGCKRLVILF | GMWDLNKNWTHSR | LAAL 1024 |
| Db | 335 | NSSLFELQMSNPL | SDSGVVELCKAL | GYDPTVLRVIL | WLGDCDVTDSGCS | SLATV 388 |

RESULT 9

US-08-910-731-8
; Sequence 8, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-08-910-731-8

Query Match 13.0%; Score 712; DB 2; Length 456;
Best Local Similarity 42.5%; Pred. No. 3.9e-61;
Matches 150; Conservative 61; Mismatches 142; Indels 0; Gaps 0;

| | | | | | | |
|----|-----|---------------------|-------------------|--------------------|-----------------------|------------|
| Qy | 671 | LLDAYSEHLAAALCTNPNI | ELSLYRNALSGRGVLL | COGLRHPNCKQLNLRKCR | ISS 730 | |
| Db | 35 | LTEEHCKDIGSALRANPS | LTCLRTNELGDAGVHL | VLQGLQSPCKIQKLSLQ | NCSLTE 94 | |
| Qy | 731 | SACEDLSAALIANKNLT | RMDSLGNVGFPGNMILL | CEGLRHPNCKQLNLRKCR | ISS 790 | |
| Db | 95 | AGCGVLPSTLRSLFT | LRHLSDNPLGDAGL | RLCEGLLDPQCHLEK | QLLEYCR | LTAASC 154 |
| Qy | 791 | QEMASVLGTNPVHVELD | ITGNALDELGLRL | COGLRHPNCKQLNLRKCR | ISS 850 | |
| Db | 155 | EPLASVLVRATRAKEL | TVSNNDIGEAGAR | VLQGLADSAQCLFTLR | ENCGLTPANCKDL 214 | |
| Qy | 851 | ASTLSVNSQSLRELD | LSNELDGLVLL | CEGLRHPNCKQLNLRKCR | ISS 910 | |
| Db | 215 | CGIVASQASLRELA | LSGNKLGDMALCP | GLLHPSRSLFTLWI | WECGITAKCGCD | LCRV 274 |
| Qy | 911 | LQAHNLRDELDSFND | LGDWGLWLAEGLO | HPACRLQKWLDS | CGLTAKACENLYFTLGI 970 | |

Db 275 LRAKESLKLKSLAGNEDGELCETLLEPGQOLESWVKSCSFTAAACPHFSVLAQ 334
Qy 971 NOTLTDLYTNALGDTGVRLCKRSLSHPGCKRLVLMFGMDLNKMTSHSLAA 1023
Db 335 NRFLELQISNNRLEDAGVRELCOGLQPGSVLRVLMWLADCDVSDSSSLAA 387

RESULT 10
US-08-910-731-2
; Sequence 2, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-910-731-2

Query Match 13.0%; Score 709; DB 2; Length 456;
Best Local Similarity 42.7%; Pred. No. 7.7e-61;
Matches 151; Conservativity 60; Mismatches 143; Indels 0; Gaps 0;
Qy 671 LLDAYSSEHAAALCTNPNLIELSLYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRISS 730
Db 35 LTBHCKDIGSALRANPSLTELCLRTNELGDAGVHLVQLGQSPCKIQKLSLQNCSLTE 94
Qy 731 SACEIDLAAIANKNLTRMDLSGNVGFPGMMLLCEGLRHPQCRQLQMIQRKCOLESAC 790
Db 95 AGCGVLPSTLRSPLTRELHLSNPLGDAGRLLCGLLDPQCHLEKQLEYCRKLTAAAC 154
Qy 791 QEWASVLGTPHVLVDLTGNALEDGLRLLCQGLRHPVCELRITLWLKICELTAAACDEL 850
Db 155 EPLASVLRATRALKELTVSNNDIGEARGVILGQGLADSAQOLETLRLNLCGLTPANCKOL 214

Qy 851 ASTLSVNQSURELDLSNELGDLGVLLLCBGLRHPTCKLOTLRIGICRLGSAACEGLSVV 910
Db 215 CGIVASQASRELDDSLASGLDAGIAELCPGLASPASRLKTLMLWECDITASGCRDLCRV 274
Qy 911 LQAHNLRLEDLSFNDLGDWGLMLLAGLOHPACRLQKWLDSGLTAKACENLYFTLGI 970
Db 275 LQAKETLKLSELNAGNLGDEGARLLCESLQPGQLESWVKSCSLTAAACQHVSLMLTQ 334
Qy 971 NOTLTDLYTNALGDTGVRLCKRSLSHPGCKRLVLMFGMDLNKMTSHSLAA 1024
Db 335 NKHLELQLSSNKLGDGSIQELCOALSOPGTLRLVLCIGDCEVTNSGCCSSLASL 388

RESULT 11
US-08-795-395-2
; Sequence 2, Application US/08795395
; Patent No. 5965399
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Cloning and Expression of Rat Liver and
; TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,395
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-795-395-2

Query Match 13.0%; Score 709; DB 2; Length 456;
Best Local Similarity 42.7%; Pred. No. 7.7e-61;
Matches 151; Conservativity 60; Mismatches 143; Indels 0; Gaps 0;
Qy 671 LLDAYSSEHAAALCTNPNLIELSLYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRISS 730
Db 35 LTBHCKDIGSALRANPSLTELCLRTNELGDAGVHLVQLGQSPCKIQKLSLQNCSLTE 94
Qy 731 SACEIDLAAIANKNLTRMDLSGNVGFPGMMLLCEGLRHPQCRQLQMIQRKCOLESAC 790
Db 95 AGCGVLPSTLRSPLTRELHLSNPLGDAGRLLCGLLDPQCHLEKQLEYCRKLTAAAC 154
Qy 791 QEWASVLGTPHVLVDLTGNALEDGLRLLCQGLRHPVCELRITLWLKICELTAAACDEL 850
Db 155 EPLASVLRATRALKELTVSNNDIGEARGVILGQGLADSAQOLETLRLNLCGLTPANCKOL 214

```

851  ASTLSVNQSLRELDLSINELGDGLVLLICEGLRHPTCKLQTLRLGICRLGSAACBGLSV 910
      :: |||||:: |||||:: |||||:: |||||:: |||||:: |||||:: |||||::
215  CGIVASQASRELDLGSNGLGDAGTAECLPGLLSPASRLKTKIWLMECDITASGCRDLCRV 274
      :: |||||:: |||||:: |||||:: |||||:: |||||:: |||||:: |||||::
911  LQANHNLRELDLSFNFDGWLIIAEGLQHPACRLQKLWLDSCLGTAKACENLYFTLGI 970
      :: |||||:: |||||:: |||||:: |||||:: |||||:: |||||:: |||||::
275  LQAKETLKELSLGNKLGDEGARLLCESLLQPGQLESILWVKVSCSLTAACCHVSLMLTQ 334
      :: |||||:: |||||:: |||||:: |||||:: |||||:: |||||:: |||||::
971  NQTLTDLYLTNNALGDTCVRLCKRLSHPGCKLRVLWLFQMDLNKQWTHRLAAL 1024
      :: |||||:: |||||:: |||||:: |||||:: |||||:: |||||:: |||||::
335  NKHLLEQLSSNKLGSIGIELCQALSQPGTTLRVLCGLDCEVTNSGCGCSLASL 388
      :: |||||:: |||||:: |||||:: |||||:: |||||:: |||||:: |||||::

RESULT 12
US-10-014-269-3
; Sequence 3, Application US/10014269
; Patent NO. 6835815
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Ogur, Yasunori
; TITLE OF INVENTION: NOD2 Nucleic Acids and Proteins
; FILE REFERENCE: UM-06645
; CURRENT APPLICATION NUMBER: US/10/014,269
; CURRENT FILING DATE: 2001-10-25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-014-269-3

```

| | | | |
|----|------|--|------|
| Qy | 632 | SAQVLIHLYGATYSADGEDRARCSCAGAHITLVQLRPERTVLDDAYSEHLAAALCTNPNIIE | 691 |
| Db | 746 | -----SVGPT-----ECAAFAVLQHLP-- | 767 |
| Qy | 692 | LSLYRNALSGRGVKLL--COGLRHPNCKLONFLKCRIRSSACEDLSAALIANKNLTRM | 749 |
| Db | 768 | LQIDYNSVGDIGVEQLPFCIGV---CKAIIYL--- | 796 |
| Qy | 750 | DLSGNGVGFFGMMLLCGLGRHPQCRLQMIOLRRKCQLESACQEMASVLGTNPHLVELDLT | 809 |
| Db | 797 | ---DNNISDRGI CKILFECALHCE-QIQKLAFNNKLTDCGAHSMAKLLACRONFLARLG | 852 |
| Qy | 810 | GNALEDIGLRLLCQGLRHVPVCRURTLWLKICRTATAACDELASTLSVNOSLRDELDLSLNE | 869 |
| Db | 853 | NNYTITAAGAQAIVLAEGRLG-----NTSLQFLGFWGNR | 883 |
| Qy | 870 | LGDLGVLLICEGI-RHPTCKLOTRLGICPLGSAAACEGLSVVLQANHNLRDELDLSFNDLG | 928 |
| Db | 884 | VGBEGAQALAEALGDHQSS--LRWLISVGNNGISVGGAQALAMLAKNVMLEELCLEENHIQ | 941 |
| Qy | 929 | DWGLWLLAEGLQHPACPRLQKLWLDSCGLTAKACENLYFTLGINQTITDLYLTNNA---- | 983 |
| Db | 942 | DEGVCSIAEGLKKNS-SLKILSLNNCIITVLGAEALLQALERNDTILEVWLRCNTFSLEE | 1000 |
| Qy | 984 | ----LGDTCVRLL | 992 |
| Db | 1001 | VDKIGCRDTELL | 1012 |

RESULT 13
 US-10-014-269-2
 ; Sequence 2, Application US/10014269
 ; Patent No. 6835815
 ; GENERAL INFORMATION:
 ; APPLICANT: Nunez, Gabriel
 ; APPLICANT: Inohara, Naohiro
 ; APPLICANT: Ogur, Yasunori
 ; TITLE OF INVENTION: MOD2 Nucleic Acids and Proteins
 ; FILE REFERENCE: UM-06645
 ; CURRENT APPLICATION NUMBER: US/10/014,269
 ; CURRENT FILING DATE: 2001-10-26
 ; NUMBER OF SEQ ID NOS: 52
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 1040
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-014-269-2

| | | | |
|----|------|--|------|
| Qy | 438 | AADGLWNOXILLPEEODLRKHGLDGEDVSFAFLNMNIFQKDI--NCERYYSVFTHLSQOEPPFA | 495 |
| Db | 554 | ALWGLMCCYYFSAQLOQAQVSPDDIS--LGFLVRKGVVPGSTAPLEFLHITPQCFFA | 611 |
| Qy | 496 | AMWYILDEG-----EGGAGPDODVTRILLTEVAFGER | 526 |
| Db | 612 | AFYLAISADVPPALLRHFLNCPGRPGNSPMARLLPTMCIQASEGKDVSSVAALLQK-----AEP | 668 |
| Qy | 527 | SFLALTGRFLGFLNEE-----TRSHLEKSLC--WKVSPHIKMDLLOWIQSKAQ | 573 |
| Db | 669 | HNLOITAAFLAGLLSREHWGLLAECOTSEKALLRRQACARWCLARSLRKH--FHSIPPAAP | 727 |
| Qy | 574 | SDGSTL--QQGSLEFFSCLYEIOREEFIOQALSHFQVIVVSNIAKMEHMYSSFLCKRCR | 631 |
| Db | 728 | GEAKSVHAMPGFIMILRSLEYEMQERLARKAARGLV-----GHLKLTFC----- | 772 |
| Qy | 632 | SAQVLHLYGATYSADGEDRACRSAGAHPLLVLQRPERTVLDDAYSEHLAAALCTNPMLIE | 691 |
| Db | 773 | -----SVGPPT-----ECAALAFVLQHLRP-----VA | 794 |
| Qy | 692 | LSLYRNALGSRGVKLL--COGLRHPNCKLQWLRLKRCRISSACEDLSAALIANKNLTRM | 749 |
| Db | 795 | LQLDYNVSGDVGVEQLLPCLVG---CKALYLR-----823 | |
| Qy | 750 | DLSCNGVGFPGPMMLLCEGLRHPQCELOMIQLRKCOLESGACQEWASVLGTNPHLVELDLT | 809 |
| Db | 824 | ---DNNISDRGICXLIICALHCE--OLOKLAFLNNKLTGDCAHSMAXKLACRQNFALRLG | 879 |
| Qy | 810 | GNALEDGLRLLCCGLRHPVCRLRTLWLKI CRLTAAACDELASTLSVNVQSLRELDLSLNE | 869 |
| Db | 880 | NNYITAAGAQVLAESLGR-----NTSLQFLGFWGNR | 910 |
| Qy | 870 | LGDLGVLLLCBGL--RHPTCKLOTLRLGICRIGLSAACGLSVVLQANHNLRELDLSFNDLG | 928 |
| Db | 911 | VGDEGAQALAEALGDHOS--LRWLSLVGNNTGSGVQAQALMLAKRNVMLLEELCELENHLQ | 968 |
| Qy | 929 | DWGLWLLAEGLOHPACRLQKWLDSCGLTAKACENLYFTLGINOTLTDLYLTNNA----- | 983 |
| Db | 969 | DEGVCSLAEGLKONS--SLKILKLNNCITYLGAALLQALERNTLILEWLRGNTFSLEE | 1027 |
| Qy | 984 | ---LGDGTGVRL | 992 |
| Db | 1028 | VDKLGCRDTRL | 1039 |

RESULT 14

US-09-949-016-9080

Sequence 9080, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 9080

LENGTH: 1043

TYPE: PRT

ORGANISM: Human

US-09-949-016-9080

| | | | |
|----|------|--|------|
| Qy | 163 | QASP--IKIETFLPEBDEERPEPRTVVVWGAAAGIGKSLAHKVMWDWADGKLFQGRPDYL | 220 |
| Db | 273 | QKSPATGLBELFSTPGHLNDDADTVLVVGEAGSGKSTLLQRLHLLWAAGQDFQ-BFLFV | 331 |
| Qy | 221 | FYINCREMNQSAQECSDMODLIFS--CWPEFSAP--LQELIRVPERLLFIIDGDEKXPSF | 276 |
| Db | 332 | FPFSCROLQCMXPLSVRTILFHCMPDVGQBDIFOLLLOHDPDRVLLTTFDGEDEFKFRF | 391 |
| Qy | 277 | HD-----POGPWMLCWEERKPTLELLNSLIRKKLJPELSJLITTRPTALEKHLRLLEHP | 330 |
| Db | 392 | TDREHCSPTD-----TSVQTLLFNLLQGNLLKNARKVTVSRPAAVSAF--LRKYI | 441 |
| Qy | 331 | R-HVEILLGFSEARKEYFYKFINAEQAGQVFNVRDNEPLFTWCFVPLVGVVVTCLCQ | 389 |
| Db | 442 | RTEFNLGXFEQGIELVLRHHEPGVADRLIRLLQETSALHGLCHLPVFSVMVSKHQE | 501 |
| Qy | 390 | QL--EGGGLLRQTSRTTAVYML-----YLLSLMQPKGAPRLOPP-----PNQRGICSL | 437 |
| Db | 502 | LLLOEGG-----SPKTTTDMYLLILQHFLUHAHTPPDASOGLGPSLLRGLPTLLHLGRL | 556 |
| Qy | 438 | AADGLMNQKILFBEQDLRKHGLDGEDVSAPFNWNIPOKOI--NCERYTSPFIHLSFOBFFA | 495 |
| Db | 557 | ALWGLGMCVYFSAQLOAAQVSPDDIS--LGLVLRAGVVPGSTAPLEFLHITTFQCFFA | 614 |
| Qy | 496 | AMYIYLDG-----EGGAGPDODVTLLTTEYAFSER | 526 |
| Db | 615 | AFYILASADVPAPLLRHLFNCGRPGNSPMARLLPTWCIOASEGKDSVAALLQK--AEP | 671 |
| Qy | 527 | SFLALTSRFLPGLNBE-----TFSHLEKSLC--WKVYSPHITKMDLLOWIOSKQ | 573 |
| Db | 672 | HNQITAAFLAGLLSREHWGLLACOTSEKALLRQACARWCILARSURKH-FHSIPPAAP | 730 |
| Qy | 574 | SDGSTL--OQGSLEFFSCLYEIOEBEFIOALSHFQIVVVSNTASKMEHVVSSFCLKRCR | 631 |
| Db | 731 | GEAKSVHMPGFIWILRSLYEMOEERLARKAARGLV-----GHLKLTFC----- | 775 |
| Qy | 632 | SAQVLHYIYATYSADGEDRARCAGAHNTLVQURPBTVLLVDAYSHEHLAAALCTNPNLIE | 691 |
| Db | 776 | -----SVGPT-----ECAALAFVLOHLRRP-----VA | 797 |
| Qy | 692 | LSLYRNALSGRVKLL--COGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRM | 749 |
| Db | 798 | LQDYNSVGDIGVEQLLPCUGV-----CKALYLR----- | 826 |
| Qy | 750 | DLSGNGVGFPGMMLCEGLRHPCQLQMIQLRRCOLESGACQEMASVLTGTPNHLVELDLT | 809 |
| Db | 827 | --DNNISDRGI CKLIECALHCE-QLOKLALFNNKLTGDCAHSMAXLLACRONFLAALRG | 882 |
| Qy | 810 | GNALIEDGLRLLCQGLRHVPVCRURTWLKICRTATAACDELASTLSVNOGSLRELDLSLNE | 869 |
| Db | 883 | NNYITAGAQAQVLAELRG-----NTSLOLGFQFWGNR | 913 |
| Qy | 870 | LGDLGVLLICEGL-RHPTCKLOTLRLGI CRLGSAACEGLSVVLQANHRLRELDLSFNDLG | 928 |
| Db | 914 | VGBEGAQALAEALGDHQSG--LRWLSLVGNNGISVGQAALALMLAKNMVLEELCEENHLQ | 971 |
| Qy | 929 | DWGLWLLIAEGLQHPACRLQKLWLDSDCGLTAKACENLYFTTIGINQTLTDTLYTNA----- | 983 |
| Db | 972 | DEGVCSLAEGLKKNS-SLKILKLSNNCITYLGAELLOALERNDTILEVWLRGNTFSLEE | 1030 |
| Qy | 984 | ---LGDGTGVRL | 992 |
| Db | 1031 | VDKLGRDTRLL | 1042 |

RESULT 15
US-10-014-269-4
; Sequence 4, Application US/10014269
; Patent No. 6835815
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro

```

; APPLICANT: Ogur, Yasunori
; TITLE OF INVENTION: NOD2 Nucleic Acids and Proteins
; FILE REFERENCE: UM-06645
; CURRENT APPLICATION NUMBER: US/10/014,269
; CURRENT FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1040
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-014-269-4

```

| | | | | | |
|-----------------------|-----|--|--------------------|-------------|--------------|
| Query Match | | 9.6% | Score 526; | DB 4; | Length 1040; |
| Best Local Similarity | | 25.4%; | Pred. No. 3.6e-42; | | |
| Matches 232; | | Conservative 117; | Mismatches 339; | Indels 224; | Gaps 33; |
| QY | 163 | QASP--IKIETLPEPDEERPEPRVTVMGAAGIGKSLAHKVMLDWADGKLFQGRFDYL | 220 | | |
| DB | 270 | QKSPATLGLSELPSTGHLNDDADTVLVGEAGSGKSTLLQRLHLLWAAGQDFQ-EFLFV | 328 | | |
| QY | 221 | FYINREMNASATSCSMDLIIFS--CWPEPSAP--LQELIRVPERLLFIIDGDELKPSF | 276 | | |
| DB | 329 | FPFSRQLQCMWAPLSVRTLLFEHCWPDVGQBDIFQLLDHPDRVLLTFDGFDEKFRF | 388 | | |
| QY | 277 | HD-----POGPWCLWEEKRPTPELLNLSIRKKLIPELSLLITRPTALEKLRHLEHP | 330 | | |
| DB | 389 | TDREHCSPTDP-----TSVQTLLFNLLQGNLLKNARKVVTSRPAVSAF--LRKYI | 438 | | |
| QY | 331 | R-HVEILGFSEABERKEYFYKFNNAEAGOVFNVRDNEPLFTMCFVPLVCVVCTCLOQ | 389 | | |
| DB | 439 | RTEFNLGFSQEGIELYLRKHHPGVADRLIRLQETSALHGLCHLPVFSWVSKCHQE | 498 | | |
| QY | 390 | QL--EGGGLLRQTSRTTAYVML-----YLSLMPQKPGAPRLOPP-----PNQGLCSL | 437 | | |
| DB | 499 | LLQEGG-----SPKTTTDMYLLTLOHFLHATPPDSASQGLGPSLIRGRLEPTLLHGLR | 553 | | |
| QY | 438 | AADGLWNQKILFEODLRKHGLDGEDVSAFLNWNIFOKDI--NCERYYSFIHLSFOEPPA | 495 | | |
| DB | 554 | ALWGLMCCVYVSAQLOAAQVSPDDIS--LGFLVRAGVVPGSTAPLEFHHITFQCFFA | 611 | | |
| QY | 496 | AMYVILDEG-----EGGAGPDQDVTRLTLEYAFSER | 526 | | |
| DB | 612 | AFYLAISADVPPALLRHFLFNGRPGNSPMARLLFTMCIQASEGKSSVAALLOK---AEP | 668 | | |
| QY | 527 | SFLATSRPLPGLNEE-----TRSHLEKSLC--WKVSPHIKMDLLQWIOSKAQ | 573 | | |
| DB | 669 | HNLIQITAAFLAGLSRBHWGLLABECOTSEKALLRHQACARWCARSIRKH-FHSIPPAAP | 727 | | |
| QY | 574 | SDGSTL--QQGSLEFFSCLYEIQEEFTQQALSHFQIVVWSNIASKMHWVSSPCLKRCR | 631 | | |
| DB | 728 | GEAKSVHAMPGFIMILRSLEYQEEERLARKAARGLV-----CHLKLTFC----- | 772 | | |
| QY | 632 | SAQVLHLYGATYSADGEDPARCSAGAHTLLVQLRPERTVLDDAYSEHAAALCTNPNLIE | 691 | | |
| DB | 773 | -----SVGPT-----ECALAFLVQLHRRP-----VA | 794 | | |
| QY | 692 | LSLYRNALSGRGVKLL--COGLRHPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTRM | 749 | | |
| DB | 795 | LQLDYNSVGDIGVEQLPCLGV-----CKALYLR----- | 823 | | |
| QY | 750 | DLSGNGVGFPGWMLLCEGLRHPQCRLOMIQLRKQESGACQEMASVLGTNPHLVELDLT | 809 | | |
| DB | 824 | ---DNNISDRGICKLIECALHCE-QLQKIALFNKKLTDGCAHSMAKULACHQNFALRLG | 879 | | |
| QY | 810 | GNALEDGLRLCOGLRHPVCRFLTWLXICRLTAACDELASTLSVNSQSLRELDLSINE | 869 | | |
| DB | 880 | NNYITAAGAAQVLAELRG-----NTSLQFLGFWGNR | 910 | | |
| QY | 870 | LGDGLVLLICEGL-RHPCTKLOTLELIGICRLGSAACEGLSVVLOQANHNLRELDLSFNDLG | 928 | | |
| DB | 911 | VDEGAQAALAEALGDHQS--LRWLSLVGNNGISVGQAQALMALAKNVMLEECLCEENHLQ | 968 | | |

```

QY 929 DWGLWLLAEGLOHPACRIQKLMWDSGGLTAKACENLYFTLGINQTLDTLYLTNNA----- 983
DB 969 DEGVCSLAEGLKONS-SLKILKLSNNCITYLGAEALLQALERNDTILEVWLRGNTFSLEE 1027
QY 984 ---LGDVTGVRLL 992
DB 1028 VDKLGCRTDRL 1039

Search completed: February 8, 2005, 14:50:49
Job time : 53 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2005, 14:54:31 ; Search time 12939 Seconds
(without alignments)
11639.138 Million cell updates/sec

Title: US-10-781-294-23
Perfect score: 3108
Sequence: 1 atgtctacgaaccgcagcag.....atttgacattgctgtcta 3108

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_btg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------------|
| 1 | 3108 | 100.0 | 3108 | 6 | AX417214 Sequence |
| 2 | 3057 | 98.4 | 3300 | 6 | AX459869 Sequence |
| 3 | 3017 | 97.1 | 3731 | 9 | AV116204 Homo sapi |
| 4 | 3001 | 96.6 | 3507 | 9 | BC028069 Homo sapi |
| 5 | 2997.8 | 96.5 | 3827 | 9 | BC028069 Homo sapi |
| 6 | 2997.8 | 96.5 | 3827 | 9 | AY095146 Homo sapi |
| 7 | 2759 | 88.8 | 3563 | 9 | AY154467 Homo sapi |
| 8 | 2644.2 | 85.1 | 3466 | 6 | AX833594 Sequence |
| 9 | 2644.2 | 85.1 | 3466 | 6 | AX833594 Sequence |
| 10 | 2470 | 79.5 | 3395 | 9 | AY116206 Homo sapi |
| 11 | 2417 | 77.8 | 3221 | 9 | AY116207 Homo sapi |
| 12 | 1703 | 54.8 | 147330 | 9 | AC008753 Homo sapi |
| 13 | 1681.8 | 54.1 | 2158 | 6 | AX575503 Sequence |
| 14 | 1616 | 52.0 | 4931 | 6 | AX684291 Sequence |
| 15 | 1511 | 48.6 | 202418 | 2 | AC150815 Callithri |
| 16 | 1256.2 | 40.4 | 2159 | 10 | AY364010 Mus muscu |
| 17 | 1036.2 | 33.3 | 22334 | 2 | AC106193 Rattus no |
| 18 | 997.8 | 32.1 | 218270 | 2 | AC079499 Mus muscu |
| 19 | 803.6 | 25.9 | 4170 | 9 | AF410477 Homo sapi |

| | | | | | |
|----|-------|------|--------|----|----------------------|
| 20 | 800.4 | 25.8 | 3102 | 6 | AX299762 Sequence |
| 21 | 800.4 | 25.8 | 3579 | 9 | AF468522 Homo sapi |
| 22 | 800.4 | 25.8 | 3857 | 6 | AX299760 Sequence |
| 23 | 800.4 | 25.8 | 3857 | 9 | AF420469 Homo sapi |
| 24 | 783.4 | 25.2 | 3102 | 10 | AY355340 Mus muscu |
| 25 | 783.4 | 25.2 | 3116 | 10 | AY337285 Mus muscu |
| 26 | 783.4 | 25.2 | 3913 | 10 | AF486632 Mus muscu |
| 27 | 783.4 | 25.2 | 3918 | 10 | AY495377 Mus muscu |
| 28 | 783.4 | 25.2 | 4022 | 10 | AY495376 Mus muscu |
| 29 | 783.2 | 25.2 | 2934 | 9 | AY422168 Homo sapi |
| 30 | 700.2 | 22.5 | 3828 | 9 | AF427617 Homo sapi |
| 31 | 698.6 | 22.5 | 3237 | 9 | AY052033 Homo sapi |
| 32 | 639 | 20.6 | 1958 | 9 | AY0511283 Homo sapi |
| 33 | 639 | 20.6 | 2524 | 6 | AX417249 Sequence |
| 34 | 637.4 | 20.5 | 2528 | 9 | AF418985 Homo sapi |
| 35 | 637.4 | 20.5 | 212426 | 9 | AC104335 Homo sapi |
| 36 | 626.6 | 20.2 | 2546 | 6 | AX780309 Sequence |
| 37 | 625.4 | 20.1 | 2546 | 9 | AK027194 Homo sapi |
| 38 | 615 | 19.8 | 1503 | 9 | AY338201 Alouatta |
| 39 | 605.4 | 19.5 | 1503 | 9 | AY338204 Cebus ape |
| 40 | 597.2 | 19.2 | 1500 | 9 | AY338202 Actus aza |
| 41 | 595.6 | 19.2 | 2494 | 9 | AF231021 Homo sapi |
| 42 | 594.2 | 19.1 | 1503 | 9 | AY338205 Saguinus |
| 43 | 592.6 | 19.1 | 1503 | 9 | AY338203 Saimiri s |
| 44 | 589.4 | 19.0 | 1521 | 9 | AY338200 Colobus g |
| 45 | 589 | 19.0 | 2213 | 10 | AY337286S3 Mus muscu |

ALIGNMENTS

RESULT 1
AX417214
LOCUS AX417214 3108 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 23 from Patent WO0226780.
ACCESSION AX417214
VERSION AX417214.1 GI:21449784
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Read,J.C., Godzik,A., Chu,Z.L., Pawlowski,K., Fiorentino,L.,
Ariza,M.E. and Stehlik,C.
TITLE Paad domain-containing polypeptides, encoding nucleic acids, and
methods of use
JOURNAL Patent: WO 0226780-A 23 04-APR-2002;
The Burnham Institute (US)
FEATURES
source Location/Qualifiers
1..3108
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..3108
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD35284.1"
/db_xref="GI:21449785"
/translation="MLRTAGDGLCRSLTYLEAEVLEKFKYLGTATELGSGKIP
WGSMEKAGPLEMAQLLITFHGPEAWRLALSTFERINRKLWEQREDLVDPQET
RDVRRKFLMEDNRNARCEPVLNTRLLLVKHSNPMQVQQQLDGTGRHGT
YLVFVINCENMQSATECSMODLIFSCWPEPSAQLQELIYRPERLLFTIDGFDLKPSP
GHQASPIKLTLPDPEPEPTVVMQGAAGIGKSLAHKVLMDWADGKLFQGRFD
ILGFSEARKEYFYKYPHNAEQGVNRYVDNEPLFTMCVPLVWVVCVLCQQQL
HCGGLRQTSRTTAVMYLLSLMQPKGAPRLQPPNQRGLCSLAADAGMYNKLIF
EODLRKHGDEGVSAFLNMNIYFQKQINCRERYSFHLSFQEFFAAYYIIDEGE
GPQDVTRLTETAFSEERSFLATSLREFLGLNEETSHLEKSLCWKSPHDKMDLLQ
WICRQKSGDSTLQOQSLEFFSCGLYEQEEFTQQALSHFQIVVSNIAKMEHVS
FCLQRKSAQVLLHYGATYSADGEDRACRAGAHLLVQLRPERTVLLDAYSEHLAA
LCVNPNIELSLNAGSGVLLCQGLRHPNCKQLNLRKRKRIRSSACEDLSAAL
IANKNLTRMDLSNGVGFPGWMLICEGLRHPQCRQLQIQLRKCQLESAGCQEMASVILG

TNPHLVLDLTGNALEDLGLRLCQGLRHPVCRRLRLWLKICRLTAACDELASTLSV
NQSIRELDLSINELGDLVLLCEGLRHPCKLQTLRLGICRLGSAACENLYSVLQAN
HNRLRLDLSINELGDLVLLCEGLRHPCKLQTLRLGICRLGSAACENLYSVLQAN
TLTDLXLTNNALGDTGVRLLCRLSHPSCKLRVLMLEFGMDLNKMTSLRLAALRVKPY
LDIGC"

ORIGIN

| | | | | | | | |
|-----------------------|-----|---|---------------|-----------|--------------|--|--|
| Query Match | | 100.0%; | Score 3108; | DB 6; | Length 3108; | | |
| Best Local Similarity | | 100.0%; | Pred. No. 0; | | | | |
| Matches 3108; | | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; | | |
| QY | 1 | ATGCTACGAACCGCAGGACGAGCGGCTCTGTGCGCTGTCCACCTACTTTGGAAGAACTC | 60 | | | | |
| DB | 1 | ATGCTACGAACCGCAGGACGAGCGGCTCTGTGCGCTGTCCACCTACTTTGGAAGAACTC | 60 | | | | |
| QY | 61 | GAGGCTGTGGAACCTGAAGAAGTTCAAGTTATACCTGGGGAACCGGACAGAGCTGGGAGAA | 120 | | | | |
| DB | 61 | GAGGCTGTGGAACCTGAAGAAGTTCAAGTTATACCTGGGGAACCGGACAGAGCTGGGAGAA | 120 | | | | |
| QY | 121 | GGCAAGATCCCTGGGGAAAGCATGGAGAAGCGGCTCCCTTGAAATGGCCCAAGCTGCTC | 180 | | | | |
| DB | 121 | GGCAAGATCCCTGGGGAAAGCATGGAGAAGCGGCTCCCTTGAAATGGCCCAAGCTGCTC | 180 | | | | |
| QY | 181 | ATCACCCACTTTCGGGCCAGAGGAGGCTTGAGGTTTGGCTCTCAGCACCTTTGAGCGGATA | 240 | | | | |
| DB | 181 | ATCACCCACTTTCGGGCCAGAGGAGGCTTGAGGTTTGGCTCTCAGCACCTTTGAGCGGATA | 240 | | | | |
| QY | 241 | AACAGGAAGACCTGTGGGAGAGGACAGAGAGGACCTGTGAGGATCCCCAGGAA | 300 | | | | |
| DB | 241 | AACAGGAAGACCTGTGGGAGAGGACAGAGAGGACCTGTGAGGATCCCCAGGAA | 300 | | | | |
| QY | 301 | ACCTACAGGACATATGTCCGACAGGAAATTCGGGCTCATGAAGACCCCAATGCGCGCTA | 360 | | | | |
| DB | 301 | ACCTACAGGACATATGTCCGACAGGAAATTCGGGCTCATGAAGACCCCAATGCGCGCTA | 360 | | | | |
| QY | 361 | GGGGAATGTCTAACTCAGCCACCGGTACACCGGCTCTCTGTGTGAAGGAGCACTCA | 420 | | | | |
| DB | 361 | GGGGAATGTCTAACTCAGCCACCGGTACACCGGCTCTCTGTGTGAAGGAGCACTCA | 420 | | | | |
| QY | 421 | AACCCCATGAGGTCAGCAGCAGCTTCTGGACACAGGCGCGGACACGCGAGGACCGTG | 480 | | | | |
| DB | 421 | AACCCCATGAGGTCAGCAGCAGCTTCTGGACACAGGCGCGGACACGCGAGGACCGTG | 480 | | | | |
| QY | 481 | GGACACAGGCTAGCCCCCATCAAGATAGAGACCTCTTTGAGCGGACAGGAGCGCCCC | 540 | | | | |
| DB | 481 | GGACACAGGCTAGCCCCCATCAAGATAGAGACCTCTTTGAGCGGACAGGAGCGCCCC | 540 | | | | |
| QY | 541 | GAGCCACCGCGCACCGTGGTCAAGATAGAGACCTCTTTGAGCGGACAGGAGCGCCCC | 600 | | | | |
| DB | 541 | GAGCCACCGCGCACCGTGGTCAAGATAGAGACCTCTTTGAGCGGACAGGAGCGCCCC | 600 | | | | |
| QY | 601 | CACAAGGTGATGTGGACTGGGCGGACGGGAAGCTCTTCCAAGGAGATTTGATTATCTC | 660 | | | | |
| DB | 601 | CACAAGGTGATGTGGACTGGGCGGACGGGAAGCTCTTCCAAGGAGATTTGATTATCTC | 660 | | | | |
| QY | 661 | TTCTCATCAACTGCGAGGAGATGAACCAAGATGCGACCGGAATGCGAGCATGCAAGACCTC | 720 | | | | |
| DB | 661 | TTCTCATCAACTGCGAGGAGATGAACCAAGATGCGACCGGAATGCGAGCATGCAAGACCTC | 720 | | | | |
| QY | 721 | ATCTTCAGCTGTGCGCTGAGCCAGCGGCGCTCTCCAGGAGCTCATCCGAGTTCCCGAG | 780 | | | | |
| DB | 721 | ATCTTCAGCTGTGCGCTGAGCCAGCGGCGCTCTCCAGGAGCTCATCCGAGTTCCCGAG | 780 | | | | |
| QY | 781 | CGCCTCTTTTTCATCATCGACGCTTCGATGAGCTCAAGCTTCTTTTCCACGATCTCTCAG | 840 | | | | |
| DB | 781 | CGCCTCTTTTTCATCATCGACGCTTCGATGAGCTCAAGCTTCTTTTCCACGATCTCTCAG | 840 | | | | |
| QY | 841 | GGACCTCTGTGCTCTCTGGAGGAGAAACGGCCACGAGGCTGTCTTTTAAACAGCTTA | 900 | | | | |
| DB | 841 | GGACCTCTGTGCTCTCTGGAGGAGAAACGGCCACGAGGCTGTCTTTTAAACAGCTTA | 900 | | | | |
| QY | 901 | ATTTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGGCCCAACGGCTTTG | 960 | | | | |

| | | | | | | | |
|----|------|---|------|--|--|--|--|
| Db | 901 | ATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATCCACACGGCCCCACGGCTTTG | 960 | | | | |
| Qy | 961 | GAGAAGCTCCACCGTCTGCTGGAGACCCCGAGGATGTGGAGATCTTGGCTTCTCTGAG | 1020 | | | | |
| Db | 961 | GAGAAGCTCCACCGTCTGCTGGAGACCCCGAGGATGTGGAGATCTTGGCTTCTCTGAG | 1020 | | | | |
| Qy | 1021 | GCAGAAAGGAAGAAATCTTACAAAGTATTTCCAAATGTCAGAGCGGGGCCCAAGTC | 1080 | | | | |
| Db | 1021 | GCAGAAAGGAAGAAATCTTACAAAGTATTTCCAAATGTCAGAGCGGGGCCCAAGTC | 1080 | | | | |
| Qy | 1081 | TTCAATTAATGAGGAGCAACAGAGCTCTCTTCAATGTCGTTGCTCCCTCTGGTGTGC | 1140 | | | | |
| Db | 1081 | TTCAATTAATGAGGAGCAACAGAGCTCTCTTCAATGTCGTTGCTCCCTCTGGTGTGC | 1140 | | | | |
| Qy | 1141 | TGGGTGCTGTACCTGCTCCAGAGCAGCTGGAGGCTGGGGGCTGTTGAGACAGACG | 1200 | | | | |
| Db | 1141 | TGGGTGCTGTACCTGCTCCAGAGCAGCTGGAGGCTGGGGGCTGTTGAGACAGACG | 1200 | | | | |
| Qy | 1201 | TCCAGGACCACTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 1260 | | | | |
| Db | 1201 | TCCAGGACCACTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 1260 | | | | |
| Qy | 1261 | GGGGCCCCGCTCTCAGGCCCCCAACACAGAGAGGTTGCTCTCTTGGCGGAGAT | 1320 | | | | |
| Db | 1261 | GGGGCCCCGCTCTCAGGCCCCCAACACAGAGAGGTTGCTCTCTTGGCGGAGAT | 1320 | | | | |
| Qy | 1321 | GGGCTCTGGAATCAGAAATCTTATTTGAGGAGGAGGCTCCGGAAGCAGCGCTAGAC | 1380 | | | | |
| Db | 1321 | GGGCTCTGGAATCAGAAATCTTATTTGAGGAGGAGGCTCCGGAAGCAGCGCTAGAC | 1380 | | | | |
| Qy | 1381 | GGGGAAGAAGCT | 1440 | | | | |
| Db | 1381 | GGGGAAGAAGCT | 1440 | | | | |
| Qy | 1441 | AGGTAATACAGCTTCT | 1500 | | | | |
| Db | 1441 | AGGTAATACAGCTTCT | 1500 | | | | |
| Qy | 1501 | CTGGAGAGGGGAGGCGGGGCGGCGGAGCCAGAGAGGCTGACAGGCTGTTGACCGAG | 1560 | | | | |
| Db | 1501 | CTGGAGAGGGGAGGCGGGGCGGCGGAGCCAGAGAGGCTGACAGGCTGTTGACCGAG | 1560 | | | | |
| Qy | 1561 | TAGCGCTTTTCTGAAAGGAGCTTCTCTGCGACTCACAGCGCTTCTCTTGTGAGCTCTCT | 1620 | | | | |
| Db | 1561 | TAGCGCTTTTCTGAAAGGAGCTTCTCTGCGACTCACAGCGCTTCTCTTGTGAGCTCTCT | 1620 | | | | |
| Qy | 1621 | AACGAGAGACCAAGGAGCCACTGGAAGAGTCTCTGCTGGAAGTCTGCGCGCACATC | 1680 | | | | |
| Db | 1621 | AACGAGAGACCAAGGAGCCACTGGAAGAGTCTCTGCTGGAAGTCTGCGCGCACATC | 1680 | | | | |
| Qy | 1681 | AGATGAGCTGTTGAGTGGATCCAAAGCAAGCTCAGAGCGGCTCCACCTGCGAG | 1740 | | | | |
| Db | 1681 | AGATGAGCTGTTGAGTGGATCCAAAGCAAGCTCAGAGCGGCTCCACCTGCGAG | 1740 | | | | |
| Qy | 1741 | CAGGCTCTCTGAGTCTTCTGAGTCTTCTGAGTCTTCTGAGTCTTCTGAGTCTTCTGAG | 1800 | | | | |
| Db | 1741 | CAGGCTCTCTGAGTCTTCTGAGTCTTCTGAGTCTTCTGAGTCTTCTGAGTCTTCTGAG | 1800 | | | | |
| Qy | 1801 | CAGGCTCTGAGCCACTTCCAGGTTGATGCTGAGTCTGAGTCTGAGTCTGAGTCTGAG | 1860 | | | | |
| Db | 1801 | CAGGCTCTGAGCCACTTCCAGGTTGATGCTGAGTCTGAGTCTGAGTCTGAGTCTGAG | 1860 | | | | |
| Qy | 1861 | ATGCT | 1920 | | | | |
| Db | 1861 | ATGCT | 1920 | | | | |
| Qy | 1921 | GGCACCTTACAGCGGAGCGGGAAGACCGCGGAGTGTCTCCGAGGAGCGCACAGCTG | 1980 | | | | |
| Db | 1921 | GGCACCTTACAGCGGAGCGGGAAGACCGCGGAGTGTCTCCGAGGAGCGCACAGCTG | 1980 | | | | |
| Qy | 1981 | TTGGTGGAGCTCAGACAGAGGAGGACCGTCTGCTGAGCGCTTACAGTGAACATCTGGCA | 2040 | | | | |
| Db | 1981 | TTGGTGGAGCTCAGACAGAGGAGGACCGTCTGCTGAGCGCTTACAGTGAACATCTGGCA | 2040 | | | | |

```
Qy 2041 GCGGCCCTGTGCACCAATCCAAACCTGATAGAGCTGTCTCTGTACCGAAATGCCCTGGC 2100
Db 2041 GCGGCCCTGTGCACCAATCCAAACCTGATAGAGCTGTCTCTGTACCGAAATGCCCTGGC 2100
Qy 2101 AGCCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACCCCACTCAAACTTCAGAAC 2160
Db 2101 AGCCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACCCCACTCAAACTTCAGAAC 2160
Qy 2161 CTGAGGCTGAAGAGGTGCGGATCTCCAGCTCAGCTCGAGGACCTCTCTGACGCTC 2220
Db 2161 CTGAGGCTGAAGAGGTGCGGATCTCTCAGCTCAGCTCGAGGACCTCTCTGACGCTC 2220
Qy 2221 ATAGCCCAATAGAAATTTGACAAGATGATCTCAGTGGCAACGGCTTGGATTCACAGC 2280
Db 2221 ATAGCCCAATAGAAATTTGACAAGATGATCTCAGTGGCAACGGCTTGGATTCACAGC 2280
Qy 2281 ATGATGCTGCTTTCGAGGGGCTCGGCATCTCCAGCTCAGCTCGAGGACCTCTCTGACGCTC 2340
Db 2281 ATGATGCTGCTTTCGAGGGGCTCGGCATCTCCAGCTCAGCTCGAGGACCTCTCTGACGCTC 2340
Qy 2341 AGGAAGTGTGAGTGGAGTCCGGGCTTGTTCAGGAGATGGCTTCTGTGCTCGGCACCAAC 2400
Db 2341 AGGAAGTGTGAGTGGAGTCCGGGCTTGTTCAGGAGATGGCTTCTGTGCTCGGCACCAAC 2400
Qy 2401 CCACATCTGTTGAGTTGGACCTGACAGGAAATGCACCTGGAGGATTTGGGCTCGAGTTA 2460
Db 2401 CCACATCTGTTGAGTTGGACCTGACAGGAAATGCACCTGGAGGATTTGGGCTCGAGTTA 2460
Qy 2461 CTATGCCAGGACTGAGCACCACCTGCTGAGAGATGAGATGAGATTTGGGCTCGAGATCTGC 2520
Db 2461 CTATGCCAGGACTGAGCACCACCTGCTGAGAGATGAGATGAGATTTGGGCTCGAGATCTGC 2520
Qy 2521 CGCCTCACTGCTGCTGCTGAGCTGAGCTGGGCTCAACTCTCAGTGTGAACACAGAGCTG 2580
Db 2521 CGCCTCACTGCTGCTGCTGAGCTGAGCTGGGCTCAACTCTCAGTGTGAACACAGAGCTG 2580
Qy 2581 AGAGAGTGAACCTGAGCTGAATGAGCTGGGACCTCGGGGTGCTGCTGCTGTGTGAG 2640
Db 2581 AGAGAGTGAACCTGAGCTGAATGAGCTGGGACCTCGGGGTGCTGCTGCTGTGTGAG 2640
Qy 2641 GGCCTCAGGACCTCCAGTGAAGCTCGAGACCTCGGCTGGGCTGAGGCTGAGGCTGGC 2700
Db 2641 GGCCTCAGGACCTCCAGTGAAGCTCGAGACCTCGGCTGGGCTGAGGCTGAGGCTGGC 2700
Qy 2701 TCTCGGCTGTGAGGCTTCTTCTGTGTGCTCCAGGCAACCAACCTCCGGGAGCTG 2760
Db 2701 TCTCGGCTGTGAGGCTTCTTCTGTGTGCTCCAGGCAACCAACCTCCGGGAGCTG 2760
Qy 2761 GACTTGAGTTTCAACGACCTGGGAGACTGGGCTGTGGTGTGCTGGCTGAGGGCTGCAA 2820
Db 2761 GACTTGAGTTTCAACGACCTGGGAGACTGGGCTGTGGTGTGCTGGCTGAGGGCTGCAA 2820
Qy 2821 CATCCGCTGCGAGCTCCAGAAATCTGTGGCTGATAGCTGTGGCTCCACAGCCAAAGCT 2880
Db 2821 CATCCGCTGCGAGCTCCAGAAATCTGTGGCTGATAGCTGTGGCTCCACAGCCAAAGCT 2880
Qy 2881 TGTGAGATCTTTACTTCACTTCCCTGGGATCAACAGACCTTGAACCGACCTTTTACTGACC 2940
Db 2881 TGTGAGATCTTTACTTCACTTCCCTGGGATCAACAGACCTTGAACCGACCTTTTACTGACC 2940
Qy 2941 AACCAACGCTTAGGGGACACAGGTGTCCGACTGTCTTTGCAAGCGGCTGAGCCATCTTGGC 3000
Db 2941 AACCAACGCTTAGGGGACACAGGTGTCCGACTGTCTTTGCAAGCGGCTGAGCCATCTTGGC 3000
Qy 3001 TGCAAACTCCAGTCTCTGTTATTTGGGATGAGCTGTGAATAAAATGACCAAGTAGG 3060
Db 3001 TGCAAACTCCAGTCTCTGTTATTTGGGATGAGCTGTGAATAAAATGACCAAGTAGG 3060
Qy 3061 TTGGCAGGCTTCAGTAACAAACCTTATTTGACATTTGCTGCTGA 3108
Db 3061 TTGGCAGGCTTCAGTAACAAACCTTATTTGACATTTGCTGCTGA 3108
```

```
RESULT 2
AX459869
LOCUS 3300 bp DNA linear PAT 30-NOV-2002
DEFINITION Sequence 12 from Patent WO0240668.
ACCESSION AX459869
VERSION AX459869.1 GI:21725643
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Tschopp, J. and Martinon, F.
TITLE Proteins and dna sequences underlying these proteins used for
treating inflammations
JOURNAL Patent: WO 0240668-A 12 23-MAY-2002;
Apotech Research and Development Ltd. (CH)
FEATURES
source
1.3300
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="NALP3/PYS.dna"
ORIGIN
Query Match 98.4%; Score 3057; DB 6; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3057; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCTACGAACCCGAGGACGAGGACGGCTCTGTGCGCTGTCCACCTACTTTGGAAGAACTC 60
Db 1 ATGCTACGAACCCGAGGACGAGGACGGCTCTGTGCGCTGTCCACCTACTTTGGAAGAACTC 60
Qy 61 GAGCTGTGAACTGGAAGAGTTCAAGTTATACCTGGGACCCGACAGAGCTGGGAGAA 120
Db 61 GAGCTGTGAACTGGAAGAGTTCAAGTTATACCTGGGACCCGACAGAGCTGGGAGAA 120
Qy 121 GGCAAGATCCCTGGGGAAGCATGGAGAGCGCGCTCCCTGGAAATGGCCAGCTGCTC 180
Db 121 GGCAAGATCCCTGGGGAAGCATGGAGAGCGCGCTCCCTGGAAATGGCCAGCTGCTC 180
Qy 181 ATCAACCACTTTGGGCGCAGAGGAGCTGTGAGGTGTGCTCTCAGCACTTTGAGCGGATA 240
Db 181 ATCAACCACTTTGGGCGCAGAGGAGCTGTGAGGTGTGCTCTCAGCACTTTGAGCGGATA 240
Qy 241 AACAGGAGACCTGTGGGAGAGGACAGAGAGGACCTGTGAGGATCCCCAGGAA 300
Db 241 AACAGGAGACCTGTGGGAGAGGACAGAGAGGACCTGTGAGGATCCCCAGGAA 300
Qy 301 ACCTACAGGACATATGTCCGACAGAAATTCGGGCTCATGGAAGACCGCAATGCGCGCTA 360
Db 301 ACCTACAGGACATATGTCCGACAGAAATTCGGGCTCATGGAAGACCGCAATGCGCGCTA 360
Qy 361 GGGGAATGTCTAACTCAGCCACCGGTACACCGGCTCTGTGTGTGGAAGGAGCACTCA 420
Db 361 GGGGAATGTCTAACTCAGCCACCGGTACACCGGCTCTGTGTGTGGAAGGAGCACTCA 420
Qy 421 AACCCCATGAGCTCAGAGAGCTTCTGGAACAGCCCGGGGACAGCGAGGACCGGTG 480
Db 421 AACCCCATGAGCTCAGAGAGCTTCTGGAACAGCCCGGGGACAGCGAGGACCGGTG 480
Qy 481 GGACACAGGCTAGCCCCATCAAGATAGAGACCTCTTTGAGCCAGACAGGAGCGCCCC 540
Db 481 GGACACAGGCTAGCCCCATCAAGATAGAGACCTCTTTGAGCCAGACAGGAGCGCCCC 540
Qy 541 GAGCCACCGCGCACCGGTGTGTCAGAGCGCGGAGGATAGGCAAGTCCATGCTGGCA 600
Db 541 GAGCCACCGCGCACCGGTGTGTCAGAGCGCGGAGGATAGGCAAGTCCATGCTGGCA 600
Qy 601 CACAAGGTGATGCTGGACTGGGCGGACGGGAAGCTCTTCCAGGACAGATTGATTATCTC 660
Db 601 CACAAGGTGATGCTGGACTGGGCGGACGGGAAGCTCTTCCAGGACAGATTGATTATCTC 660
```

| | | | | | | | |
|----|------|---|------|----|------|---|------|
| Qy | 661 | TTCTACATCACTGCGAGGAGATGAACACGAGTGCACCGAATGCAGCATGCAGACCTC | 720 | Qy | 1741 | CAGGCTCTCTGGAGTCTCTCAGCTGCTTGTACGAGATCCAGGAGGAGGATTTATCCAG | 1800 |
| Db | 661 | TTCTACATCACTGCGAGGAGATGAACACGAGTGCACCGAATGCAGCATGCAGACCTC | 720 | Db | 1741 | CAGGCTCTCTGGAGTCTCTCAGCTGCTTGTACGAGATCCAGGAGGAGGATTTATCCAG | 1800 |
| Qy | 721 | ATCTTCAGCTGCTGGCTGAGCCAGCGCCCTCTCCAGGAGCTCATCCGAGTTCCCGAG | 780 | Qy | 1801 | CAGGCCCTGAGCACTTCAGGTGATCGTGGTTCAGCAACATTTGCCCTCAAGATGAGGAC | 1860 |
| Db | 721 | ATCTTCAGCTGCTGGCTGAGCCAGCGCCCTCTCCAGGAGCTCATCCGAGTTCCCGAG | 780 | Db | 1801 | CAGGCCCTGAGCACTTCAGGTGATCGTGGTTCAGCAACATTTGCCCTCAAGATGAGGAC | 1860 |
| Qy | 781 | CGCTCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAGCTCTTCTTTTCCAGCATCTCAG | 840 | Qy | 1861 | ATGGTCTCTCTGTTCTGATGCGCTGCAGGAGCGCCAGGCTCTGCATCTGTATGGC | 1920 |
| Db | 781 | CGCTCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAGCTCTTCTTTTCCAGCATCTCAG | 840 | Db | 1861 | ATGGTCTCTCTGTTCTGATGCGCTGCAGGAGCGCCAGGCTCTGCATCTGTATGGC | 1920 |
| Qy | 841 | GGACCTCTGCTCTCTCTGGAGGAGAAACGGCCCAACGAGTCTCTTTTAAACGCTTA | 900 | Qy | 1921 | GCACCTACAGCGGAGCGGGAACACCGCGAGGTCTCCGAGGAGCGCACGCTG | 1980 |
| Db | 841 | GGACCTCTGCTCTCTCTGGAGGAGAAACGGCCCAACGAGTCTCTTTTAAACGCTTA | 900 | Db | 1921 | GCACCTACAGCGGAGCGGGAACACCGCGAGGTCTCCGAGGAGCGCACGCTG | 1980 |
| Qy | 901 | ATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATACCAACAGGGCCACGGCTTTG | 960 | Qy | 1981 | TTGTGTCAGCTCAGACCAAGAGGACCGTTCTGTGTCGAGCGCTACAGTGAACATCTGGCA | 2040 |
| Db | 901 | ATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATACCAACAGGGCCACGGCTTTG | 960 | Db | 1981 | TTGTGTCAGCTCAGACCAAGAGGACCGTTCTGTGTCGAGCGCTACAGTGAACATCTGGCA | 2040 |
| Qy | 961 | GAGAAGCTCCAACCTGCTGGAGCACCCAGGATGTGAGATCTCTGGCTCTCTGAG | 1020 | Qy | 2041 | CGGGCCCTGTGCACCAATCCAAACCTGATAGAGCTGTCTGTACCGAAATGCCCTGGGC | 2100 |
| Db | 961 | GAGAAGCTCCAACCTGCTGGAGCACCCAGGATGTGAGATCTCTGGCTCTCTGAG | 1020 | Db | 2041 | CGGGCCCTGTGCACCAATCCAAACCTGATAGAGCTGTCTGTACCGAAATGCCCTGGGC | 2100 |
| Qy | 1021 | GCAGAAAGGAAGATCTTCTACAGTATTTCCACAATGCAGAGCGGGGCCAAGTC | 1080 | Qy | 2101 | AGCCGGGGGTGAAGCTGCTCTGTCAAGCACTCAGACACCCCACTGCAGAAC | 2160 |
| Db | 1021 | GCAGAAAGGAAGATCTTCTACAGTATTTCCACAATGCAGAGCGGGGCCAAGTC | 1080 | Db | 2101 | AGCCGGGGGTGAAGCTGCTCTGTCAAGCACTCAGACACCCCACTGCAGAAC | 2160 |
| Qy | 1081 | TTCAATTTACGTGAGGACCAAGCTCTCTTCAACATGTCTTCTGCTCCCTGGTGTGC | 1140 | Qy | 2161 | CTGAGCTGAAGAGGTGCCGATCTCCAGCTCAGCTCGAGGACCTCTCTCGAGCTCTC | 2220 |
| Db | 1081 | TTCAATTTACGTGAGGACCAAGCTCTCTTCAACATGTCTTCTGCTCCCTGGTGTGC | 1140 | Db | 2161 | CTGAGCTGAAGAGGTGCCGATCTCCAGCTCAGCTCGAGGACCTCTCTCGAGCTCTC | 2220 |
| Qy | 1141 | TGGGTGTTGTATCCTGCTCCAGCAGCAGCTGAGGGGTGGGGCTGTGTGACAGAGC | 1200 | Qy | 2221 | ATAGCCCAATAAGAAATTTGACAAAGATGGATCTCAGTGGCAACCGGGTTGGAATCCAGGC | 2280 |
| Db | 1141 | TGGGTGTTGTATCCTGCTCCAGCAGCAGCTGAGGGGTGGGGCTGTGTGACAGAGC | 1200 | Db | 2221 | ATAGCCCAATAAGAAATTTGACAAAGATGGATCTCAGTGGCAACCGGGTTGGAATCCAGGC | 2280 |
| Qy | 1201 | TCCAGGACCACTGAGTGTACATGCTCTACCTGCTGAGTCTGATGCAACCAAGCCG | 1260 | Qy | 2281 | ATGATGCTCTTTTGGAGGGCTCGGGCATTCGCCAGTCCCGAGTGCAGATGATTCAGTTG | 2340 |
| Db | 1201 | TCCAGGACCACTGAGTGTACATGCTCTACCTGCTGAGTCTGATGCAACCAAGCCG | 1260 | Db | 2281 | ATGATGCTCTTTTGGAGGGCTCGGGCATTCGCCAGTCCCGAGTGCAGATGATTCAGTTG | 2340 |
| Qy | 1261 | GGGGCCCCCGCTCCAGCCCCCACCACAGAGAGGTTGTCTCTTTGGGGGAGAT | 1320 | Qy | 2341 | AGGAAGTGCAGTGGAGTCCGGGCTTGTACAGGATGGCTTCTGTGCTCCGCAAC | 2400 |
| Db | 1261 | GGGGCCCCCGCTCCAGCCCCCACCACAGAGAGGTTGTCTCTTTGGGGGAGAT | 1320 | Db | 2341 | AGGAAGTGCAGTGGAGTCCGGGCTTGTACAGGATGGCTTCTGTGCTCCGCAAC | 2400 |
| Qy | 1321 | GGGCTCTGGAATCAGAAATCCTATTTGAGAGCAGACCTCCGGAAGCAGGCTAGAC | 1380 | Qy | 2401 | CCACATCTGTTGAGTTGCACTGCAGAGAAATGCACTGGAGATTTGGGCTGAGGTTA | 2460 |
| Db | 1321 | GGGCTCTGGAATCAGAAATCCTATTTGAGAGCAGACCTCCGGAAGCAGGCTAGAC | 1380 | Db | 2401 | CCACATCTGTTGAGTTGCACTGCAGAGAAATGCACTGGAGATTTGGGCTGAGGTTA | 2460 |
| Qy | 1381 | GGGAAGACGCTCTGCTCCTCAACATGAACATCTTCCAGAGGACATCAACTGTGAG | 1440 | Qy | 2461 | CTATGCCAGGAGCTGAGGACCCAGTCTGCAGACTACGCACTTGTGGCTGAGATCTGC | 2520 |
| Db | 1381 | GGGAAGACGCTCTGCTCCTCAACATGAACATCTTCCAGAGGACATCAACTGTGAG | 1440 | Db | 2461 | CTATGCCAGGAGCTGAGGACCCAGTCTGCAGACTACGCACTTGTGGCTGAGATCTGC | 2520 |
| Qy | 1441 | AGGTACTACAGCTTCACTTCACTGTTTCCAGAAATCTTTGAGTATGTATATATC | 1500 | Qy | 2521 | CGCTCTCAGCTGCTGCTGTCAGAGCTGGCTCAACTCTCAGTGTGAACACAGAGCTG | 2580 |
| Db | 1441 | AGGTACTACAGCTTCACTTCACTGTTTCCAGAAATCTTTGAGTATGTATATATC | 1500 | Db | 2521 | CGCTCTCAGCTGCTGCTGTCAGAGCTGGCTCAACTCTCAGTGTGAACACAGAGCTG | 2580 |
| Qy | 1501 | CTGACGAGGGGAGGGCGGGCAGGCCAGACAGGACGTCACAGGCTGTTCACCGAG | 1560 | Qy | 2581 | AGAGAGCTGACCTGAGCCCTGATGAGCTGGGGACCTCGGGGTGCTGTCTGTGTGAG | 2640 |
| Db | 1501 | CTGACGAGGGGAGGGCGGGCAGGCCAGACAGGACGTCACAGGCTGTTCACCGAG | 1560 | Db | 2581 | AGAGAGCTGACCTGAGCCCTGATGAGCTGGGGACCTCGGGGTGCTGTCTGTGTGAG | 2640 |
| Qy | 1561 | TACGGCTTTTCTGAAAGGAGCTTCTGGCACTCAACAGCCGCTTCTGTTGGACTCCTG | 1620 | Qy | 2641 | GGCTCAGGCATCCACGCTGCAAGCTCCAGACCTCGGGTGGGCATCTGCGGCTGGGC | 2700 |
| Db | 1561 | TACGGCTTTTCTGAAAGGAGCTTCTGGCACTCAACAGCCGCTTCTGTTGGACTCCTG | 1620 | Db | 2641 | GGCTCAGGCATCCACGCTGCAAGCTCCAGACCTCGGGTGGGCATCTGCGGCTGGGC | 2700 |
| Qy | 1621 | AACGAGGAGACAGGAGGAGGAGTCTCTGCTGGAAGTCTCTGCTGGAAGTCTCCCGCAG | 1680 | Qy | 2701 | TCGCGGCTGTCAGGCTCTTTCTGTGTCCTCAGGCAACCAACCACTCCGCGAGCTG | 2760 |
| Db | 1621 | AACGAGGAGACAGGAGGAGGAGTCTCTGCTGGAAGTCTCTGCTGGAAGTCTCCCGCAG | 1680 | Db | 2701 | TCGCGGCTGTCAGGCTCTTTCTGTGTCCTCAGGCAACCAACCACTCCGCGAGCTG | 2760 |
| Qy | 1681 | RAGATGACCTGTCAGTGTATCCAAAGCAAGCTCAGAGCGGCTCCACCTGTCAG | 1740 | Qy | 2761 | GACTTGAGTTTCAACGACCTTGGGAGCTGGGGCTGTGGTGTGTGGCTGAGGGGCTGCAA | 2820 |
| Db | 1681 | RAGATGACCTGTCAGTGTATCCAAAGCAAGCTCAGAGCGGCTCCACCTGTCAG | 1740 | Db | 2761 | GACTTGAGTTTCAACGACCTTGGGAGCTGGGGCTGTGGTGTGTGGCTGAGGGGCTGCAA | 2820 |
| | | | | Qy | 2821 | CATCCCGCTGTCAGACTCCAGAAACCTGTGCTGATAGCTGTGGCTCTCACAGCAAGCT | 2880 |


```
|||||
2821 CA|CCCGCTGCGAGACTCCAGAACTGTGGTGGATAGCTGTGGCTCACAGCCAAAGGCT 2880
QY TGTGAGAACTTTTACTTCACCTCGGGATCAACAGACCTTGGACCGACTTTTACTCTGAC 2940
Db TGTGAGAACTTTTACTTCACCTCGGGATCAACAGACCTTGGACCGACTTTTACTCTGAC 2940
QY AACAAAGCCCTAGGGGACACAGGTGTCCGACTGCTTTTGAAGCGGCTGAGCATCTGGC 3000
Db AACAAAGCCCTAGGGGACACAGGTGTCCGACTGCTTTTGAAGCGGCTGAGCATCTGGC 3000
QY TGCAAACTCCAGTCTCTGTTTATTTGGGATGACCTTGAATATAAATGACCCACAGT 3057
Db TGCAAACTCCAGTCTCTGTTTATTTGGGATGACCTTGAATATAAATGACCCACAGT 3057

RESULT 3
AY116204 3731 bp mRNA linear PRI 21-MAY-2003
LOCUS Homo sapiens monarch-1 mRNA, complete cds; alternatively spliced.
DEFINITION AY116204
ACCESSION AY116204
VERSION AY116204.1 GI:21711820
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3731)
AUTHORS Williams,K.L., Taxman,D.J., Linhoff,M.W., Reed,W. and Ting,J.P.Y.
TITLE Monarch-1: A Pyrin/Nucleotide-Binding Domain/Leucine-Rich Repeat
JOURNAL Protein That Controls Classical and Nonclassical MHC Class I Genes
PUBMED J. Immunol. 170 (11), 5354-5358 (2003)
12759408
REFERENCE 2 (bases 1 to 3731)
AUTHORS Williams,K.L., Linhoff,M.W., Harton,J.A. and Ting,J.P.Y.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm
PUBMED Road, Chapel Hill, NC 27599, USA
FEATURES
1. 3731
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
/cell_line="U937"
221..3409
/note="alternatively spliced"
/codon_start=1
/product="monarch-1"
/protein_id="AA075142.1"
/db_xref="GI:21711821"
/translation="MLRTAGDGLCRSLTVLEAEVLEAVELKKFLYLGTATLGECKLP
WGSMEKPLSMAQLLTHFPEAEWRLALSTFERINKDLWERGQREDLVRDTPPG
PSSLGNSTCLLEVLSTPDRPOETDYVRRKFLMEDRNALRGCSVNLSHRYRL
LLVKEHNPVQQLLDGTGHARTVHQAPIKIETLPEDEERPEPRTVMQQA
AGIKRSLHAKLMDWADKLFQGRFDYLFYINCRMNQSAECMODLIFSCWPEPS
APIQELIRPERLLFIDGDELKPSFHDPOGFWCLWEEKRPTELLNSLIRKLLP
ELSLITRTPALSKHLRHLRPHRVELGFSEERKEYFYKYPHNAEQGVNRYR
DNEFLTFVPLVWVCTCLOQLEGGLRQTSRTTAVMYLYLISLMQPKGAP
RLQPPNQRGLCSLAWQKILFEEQDLRGLDGEDVSFAFLNNIIFQKINCR
YSPITHLSFQFFAAVYILDGEAGGAPDQDVRLLTEYAFSERSFALTSRPLFGL
LNESTRHLEKSLCWKYPSPHIMDLQWIOSKQSDGSLTQOQSLEFSCLYEIOBE
FIQALSHFOVIVVSNLASMEHNVSSFLKRCRCSAQLHLVLYGATYSDGEDRACSA
GAHTLLVQLRPERVLLDAYSEHLAAALCTNPNLIELSLRYNALGSRGKLLCOGLRH
PNCQLNLRKRRISSACDLSAALIANKNTLMDLSNGVGFPGMMLLCEGRHP
QCRQLMQLRQCOLESACQEMASVLGTNPHLVELDTGNALDGLRLLCQGLRHV
CRLRLTLKI CRLTAACDELATLSVNSQSLRELDLSINELGDLGLLCEGLRHPTC
KLOTLRLGICRLGSAACEGLSVLQANHLRELDLSFNLDGDLGLWLAEGLOHPACK
LQKLWLDSCGILTAKACENLYFTLGINQTLTLDLYLTNNALGDTGTVRLLCRLSHPGCKL
KVLWLFQMDLNMKTHSLAURVTKPYLDIGC"
/region="Region: P-loop"
```

misc_feature

869..892

/note="Region: P-loop"

```
misc_feature 2069..3406
/region="Region: leucine rich repeat"
misc_feature 221..289
/region="Region: pyrin domain"
ORIGIN
Query Match 97.1%; Score 3017; DB 9; Length 3731;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 3108; Conservative 0; Mismatches 0; Indels 81; Gaps 1;
QY 1 ATGCTAGAACCGCAGGACGAGCGGCTCTGTGCGCTGTCTCAGCTACTTGGAAAGACTC 60
Db 221 ATGCTAGAACCGCAGGACGAGCGGCTCTGTGCGCTGTCTCAGCTACTTGGAAAGACTC 280
QY 61 GAGGCTGTGGAACCTGAAGAAAGTTCAAGTTATACCTGGGGACCGGACAGAGCTGGGAGAA 120
Db 281 GAGGCTGTGGAACCTGAAGAAAGTTCAAGTTATACCTGGGGACCGGACAGAGCTGGGAGAA 340
QY 121 GGCAGAGATCCCTGGGGAGCATGGAGAGCGGCTCCCTTGGAAATGGCCAGCTGCTC 180
Db 341 GGCAGAGATCCCTGGGGAGCATGGAGAGCGGCTCCCTTGGAAATGGCCAGCTGCTC 400
QY 181 ATCACCACACTTCGGGCCAGAGAGGCGCTGGAGGTTGGCTCTCAGCACCTTTGAGGGGATA 240
Db 401 ATCACCACACTTCGGGCCAGAGAGGCGCTGGAGGTTGGCTCTCAGCACCTTTGAGGGGATA 460
QY 241 AACAGGAAGACCTGTGGGAGAGAGGACAGAGAGGACCTTGGTGAGG- 288
Db 461 AACAGGAAGACCTGTGGGAGAGAGGACAGAGAGGACCTTGGTGAGGAGATACCCACCT 520
QY 289 ----- 288
Db 521 GGTGGCCGCTCCTCAGTTGGGAAACAGTCAAGTCAAGTCTCTCTTGTGCACT 580
QY 289 -----GATCCCCAGGAAACCTACAGGAGCTATGTCCGACAGAAATTCGGGCTCATG 339
Db 581 CCAAGAAAAGATCCCCAGGAAACCTACAGGAGCTATGTCCGACAGGAAATTCGGGCTCATG 640
QY 340 GAAGACCGCAATGCGCCCTAGGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTC 399
Db 641 GAAGACCGCAATGCGCCCTAGGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTC 700
QY 400 CTGCTGTGTAAGGAGCACTCAAAACCCCATGCAGGTCCAGCAGCAGCTTCTGGAACACAGGC 459
Db 701 CTGCTGTGTAAGGAGCACTCAAAACCCCATGCAGGTCCAGCAGCAGCTTCTGGAACACAGGC 760
QY 460 CGGGGACACCGGAGGACCTGGGACACACAGGCTAGCCCTCATAGATAGAGACCTCTTT 519
Db 761 CGGGGACACCGGAGGACCTGGGACACACAGGCTAGCCCTCATAGATAGAGACCTCTTT 820
QY 520 GAGCCAGACGAGGAGCGCCCGGACCGACCGCTGGTTCATGCAAGGCGCGGACGGG 579
Db 821 GAGCCAGACGAGGAGCGCCCGGACCGACCGCTGGTTCATGCAAGGCGCGGACGGG 880
QY 580 ATAGGCAAGTCCATGTGGCACAACAAGGTGATGTGGAATGGGCGGACGGGAAGCTTTC 639
Db 881 ATAGGCAAGTCCATGTGGCACAACAAGGTGATGTGGAATGGGCGGACGGGAAGCTTTC 940
QY 640 CAAGGCAAGTTCATCTTCTTCTATCACTGCGAGGAGATGAACAGAGTGCACG 699
Db 941 CAAGGCAAGTTCATCTTCTTCTATCACTGCGAGGAGATGAACAGAGTGCACG 1000
QY 700 GAATGACGATGCAAGACCTCATCTTCAGCTGTGGCTCAGCCAGCGGCTCTCCAG 759
Db 1001 GAATGACGATGCAAGACCTCATCTTCAGCTGTGGCTCAGCCAGCGGCTCTCCAG 1060
QY 760 GAGCTCATCCGAGTTCCTCGAGCGCTCTCTTTTTCATCATCGAGCTTCGATGAGCTCAAG 819
Db 1061 GAGCTCATCCGAGTTCCTCGAGCGCTCTCTTTTTCATCATCGAGCTTCGATGAGCTCAAG 1120
QY 820 CTTTCTTTTCACGATTCCTCAGGAGCCCTGTGTCCTCTGCTGGGAGGAGAAACGGCCACG 879
Db 1121 CTTTCTTTTCACGATTCCTCAGGAGCCCTGTGTCCTCTGCTGGGAGGAGAAACGGCCACG 1180
```


| | | | | | | | |
|----|------|--|------|----|------|--|------|
| Qy | 880 | GAGCTGCTTCTTAAACAGCTTAAATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATC | 939 | Qy | 1960 | TCCGACGAGGCGCACACGCTGTTGGTGAGCTCAGACCCAGAGAGACCGTTCTGCTGAC | 2019 |
| Db | 1181 | GAGCTGCTTCTTAAACAGCTTAAATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATC | 1240 | Db | 2261 | TCCGACGAGGCGCACACGCTGTTGGTGAGCTCAGACCCAGAGAGACCGTTCTGCTGAC | 2320 |
| Qy | 940 | ACCACAGGCGCCACGGCTTTGGAGAGAGCTCCACCGTCTGCTGAGACACCCAGGCAATGG | 999 | Qy | 2020 | GCTTACAGTGAACATCTGGCAGCGGCTCTGTGCACCAATTCAAACCTGATAGAGCTGCT | 2079 |
| Db | 1241 | ACCACAGGCGCCACGGCTTTGGAGAGAGCTCCACCGTCTGCTGAGACACCCAGGCAATGG | 1300 | Db | 2321 | GCTTACAGTGAACATCTGGCAGCGGCTCTGTGCACCAATTCAAACCTGATAGAGCTGCT | 2380 |
| Qy | 1000 | GAGATCCTGGCTTCTCTGAGGACAGAAAGGAAGAAATFACTTTCTCAAGTATTTCCACAAT | 1059 | Qy | 2080 | CTGTACCGAATGCCCTGGGACGCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACAC | 2139 |
| Db | 1301 | GAGATCCTGGCTTCTCTGAGGACAGAAAGGAAGAAATFACTTTCTCAAGTATTTCCACAAT | 1360 | Db | 2381 | CTGTACCGAATGCCCTGGGACGCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACAC | 2440 |
| Qy | 1060 | GCAGAGCAGGCGGCCAAGTCTTCAATTAAGTAGAGGAACAAAGAGCTCTCTTCCACATG | 1119 | Qy | 2140 | CCCAACTGCAAATTCAGAACTGAGGCTGAAGAGGTGCCGATCTCCAGCTCAGCCTGC | 2199 |
| Db | 1361 | GCAGAGCAGGCGGCCAAGTCTTCAATTAAGTAGAGGAACAAAGAGCTCTCTTCCACATG | 1420 | Db | 2441 | CCCAACTGCNAATTCAGAACTGAGGCTGAAGAGGTGCCGATCTCCAGCTCAGCCTGC | 2500 |
| Qy | 1120 | TGCTTCCTCCCTGGTGTGCTGGTGTGTGTACCTGCCCTCCAGCAGCAGCTGGAGGGT | 1179 | Qy | 2200 | GAGGACCTCTCTGACGCTCTCATAGCCTCATAGCCAAATAGAAATTTGACAGGATGATCTCAGTGC | 2259 |
| Db | 1421 | TGCTTCCTCCCTGGTGTGCTGGTGTGTGTACCTGCCCTCCAGCAGCAGCTGGAGGGT | 1480 | Db | 2501 | GAGGACCTCTCTGACGCTCTCATAGCCAAATAGAAATTTGACAGGATGATCTCAGTGC | 2560 |
| Qy | 1180 | GGGGGCTGTGTGACAGAGCTCCAGGACCAACCACTGACGTGTACATGCTCTACTGCTG | 1239 | Qy | 2260 | AACGGGTTGGAATCCCAAGGCTATGCTCTTCCGAGGGCTCGGGCATCCCAAGTGC | 2319 |
| Db | 1481 | GGGGGCTGTGTGACAGAGCTCCAGGACCAACCACTGACGTGTACATGCTCTACTGCTG | 1540 | Db | 2561 | AACGGGTTGGAATCCCAAGGCTATGCTCTTCCGAGGGCTCGGGCATCCCAAGTGC | 2620 |
| Qy | 1240 | AGTCTGATGMAACCCAAAGCGGGGGCCCGGCTCCAGCCCCCAACCAACGAGAGGG | 1299 | Qy | 2320 | AGGCTGCAGATGATTCAGTTGAGGAAGTGTCAAGTGGAGTCCGGGCTTGTCAAGAGATG | 2379 |
| Db | 1541 | AGTCTGATGMAACCCAAAGCGGGGGCCCGGCTCCAGCCCCCAACCAACGAGAGGG | 1600 | Db | 2621 | AGGCTGCAGATGATTCAGTTGAGGAAGTGTCAAGTGGAGTCCGGGCTTGTCAAGAGATG | 2680 |
| Qy | 1300 | TTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAATCTCTATTGAGGAGCAGGAC | 1359 | Qy | 2380 | GCTTCTGTCTCGGCACCAACCCACATCTGCTGTGAGTTGGACCTGACAGAAATGCACATG | 2439 |
| Db | 1601 | TTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAATCTCTATTGAGGAGCAGGAC | 1660 | Db | 2681 | GCTTCTGTCTCGGCACCAACCCACATCTGCTGTGAGTTGGACCTGACAGAAATGCACATG | 2740 |
| Qy | 1360 | CTCCGGAAGCAGCGCTAGACGGGGAAGCGTCTCTGCTTCTCTCAATGAACATCTTC | 1419 | Qy | 2440 | GAGGATTTGGGCTGAGGTTACTATGCCAGGAGCTGAGGCAACCCAGTCTGACAGCTACCG | 2499 |
| Db | 1661 | CTCCGGAAGCAGCGCTAGACGGGGAAGCGTCTCTGCTTCTCTCAATGAACATCTTC | 1720 | Db | 2741 | GAGGATTTGGGCTGAGGTTACTATGCCAGGAGCTGAGGCAACCCAGTCTGACAGCTACCG | 2800 |
| Qy | 1420 | CAGAAGGACATCACTGTGAGAGTACTACAGCTTCTCACTGAGTTTCCAGGAATTC | 1479 | Qy | 2500 | ACTTTGTGCTGAAGATCTGCCGCTCACTGCTGTGCTGTGACGAGCTGSCCTCAACT | 2559 |
| Db | 1721 | CAGAAGGACATCACTGTGAGAGTACTACAGCTTCTCACTGAGTTTCCAGGAATTC | 1780 | Db | 2801 | ACTTTGTGCTGAAGATCTGCCGCTCACTGCTGTGCTGTGACGAGCTGSCCTCAACT | 2860 |
| Qy | 1480 | TTTGACGCTATGTAATCTCTGGAACGAGGGGAGCGGGGAGGCGGAGCCAGACAGGAC | 1539 | Qy | 2560 | CTCAGTGTGACACGAGGCTGAGAGCTGAGGCTGAGGCTGAAATGAGCTGGGGGACCTC | 2619 |
| Db | 1781 | TTTGACGCTATGTAATCTCTGGAACGAGGGGAGCGGGGAGGCGGAGCCAGACAGGAC | 1840 | Db | 2861 | CTCAGTGTGACACGAGGCTGAGAGCTGAGGCTGAGGCTGAAATGAGCTGGGGGACCTC | 2920 |
| Qy | 1540 | GTGACAGGCTGTGACCGAGTACCGGTTTCTGAAAGGAGCTTCTGGGACTCACCAGC | 1599 | Qy | 2620 | GGGGTGTGCTGTGTGTGAGGGCTCAGGCTTCCACGCTGCAAGCTCCAGACCTTCGCG | 2679 |
| Db | 1841 | GTGACAGGCTGTGACCGAGTACCGGTTTCTGAAAGGAGCTTCTGGGACTCACCAGC | 1900 | Db | 2921 | GGGGTGTGCTGTGTGTGAGGGCTCAGGCTTCCACGCTGCAAGCTCCAGACCTTCGCG | 2980 |
| Qy | 1600 | CGCTTCTGTGAGTCTCTGAACGAGGAGACGAGGACCACTGGGAGAGGCTCTCTGC | 1659 | Qy | 2680 | TTGGGCACTGCGGCTGGGCTCTGCCGCTGTGAGGGTCTTTCTGCTGTGCTCCAGGGCC | 2739 |
| Db | 1901 | CGCTTCTGTGAGTCTCTGAACGAGGAGACGAGGACCACTGGGAGAGGCTCTCTGC | 1960 | Db | 2981 | TTGGGCACTGCGGCTGGGCTCTGCCGCTGTGAGGGTCTTTCTGCTGTGCTCCAGGGCC | 3040 |
| Qy | 1660 | TGGAAGGTCTGCGCCACATCAAGATGACCTGTTGACAGTGGATCCAAAGCAAAAGCTCAG | 1719 | Qy | 2740 | AACCAACCTCCGAGGCTGGAATTTGAGTTTCAACGACTGGGAGACTGGGGCTCTGTGG | 2799 |
| Db | 1961 | TGGAAGGTCTGCGCCACATCAAGATGACCTGTTGACAGTGGATCCAAAGCAAAAGCTCAG | 2020 | Db | 3041 | AACCAACCTCCGAGGCTGGAATTTGAGTTTCAACGACTGGGAGACTGGGGCTCTGTGG | 3100 |
| Qy | 1720 | AGCGACGGCTCCACCTCGAGCAGGGCTCTTGGAGTCTTCAAGTGTCTGTAGAGATC | 1779 | Qy | 2800 | TTGCTGTGCTGAGGGGCTGCAACATCCCGCTCGAGACTCCAGAAAATGTGGCTGGATAGC | 2859 |
| Db | 2021 | AGCGACGGCTCCACCTCGAGCAGGGCTCTTGGAGTCTTCAAGTGTCTGTAGAGATC | 2080 | Db | 3101 | TTGCTGTGCTGAGGGGCTGCAACATCCCGCTCGAGACTCCAGAAAATGTGGCTGGATAGC | 3160 |
| Qy | 1780 | CAGGAGGAGGAGTTTATCCAGCAGGCTTGAGGCACTTCCAGGTGATCGTGGTCAAGAAC | 1839 | Qy | 2860 | TGTGGCTTCAACGCAAGGCTTGTGAGATCTTTATCTTACCTCGGGATCAACAGAGACC | 2919 |
| Db | 2081 | CAGGAGGAGGAGTTTATCCAGCAGGCTTGAGGCACTTCCAGGTGATCGTGGTCAAGAAC | 2140 | Db | 3161 | TGTGGCTTCAACGCAAGGCTTGTGAGATCTTTATCTTACCTCGGGATCAACAGAGACC | 3220 |
| Qy | 1840 | ATTGCTCCCAAGATGGAGCAGATGTCTCTGTTCTGTGTGAAGCGCTGACGAGCGCC | 1899 | Qy | 2920 | TTGACCGACCTTTTACCTGACCAACGCTTATAGGGGACACAGGTGTCCGACTGCTTTC | 2979 |
| Db | 2141 | ATTGCTCCCAAGATGGAGCAGATGTCTCTGTTCTGTGTGAAGCGCTGACGAGCGCC | 2200 | Db | 3221 | TTGACCGACCTTTTACCTGACCAACGCTTATAGGGGACACAGGTGTCCGACTGCTTTC | 3280 |
| Qy | 1900 | CAGGTGCTGACTTGTATGGCCACCTACAGCGGAGCGGGAGAGACCGCGGAGGTGC | 1959 | Qy | 2980 | AAGCGGCTGAGCCATCTGCTGCAAACTCCGAGTCTCTGTTTATTTGGGATGAGCTG | 3039 |
| Db | 2201 | CAGGTGCTGACTTGTATGGCCACCTACAGCGGAGCGGGAGAGACCGCGGAGGTGC | 2260 | Db | 3281 | AAGCGGCTGAGCCATCTGCTGCAAACTCCGAGTCTCTGTTTATTTGGGATGAGCTG | 3340 |
| | | | | Qy | 3040 | AATAAATGACCCACAGTAGGTTGGGAGCGCTTCGAGTAACAAACCTTATTGGACATT | 3099 |

| | | | | |
|------------|---|--|------|--|
| Db | 3341 | AATAAATGACCCAGTAGGTGGCAGCGCTTCAGTAACAAAACCTTATTGGCATT 3400 | | |
| Qy | 3100 | GGCTGCTGA | 3108 | |
| Db | 3401 | GGCTGCTGA | 3409 | |
| RESULT 4 | BC028069 | LOCUS | | |
| DEFINITION | Homo sapiens NACHT, leucine rich repeat and PYD containing 12, mRNA (CDNA clone MGC:40117 IMAGE:5212737), complete cds. | | | |
| ACCESSION | BC028069 | | | |
| VERSION | BC028069.1 | GI:20380399 | | |
| KEYWORDS | MGC. | | | |
| SOURCE | Homo sapiens (human) | | | |
| ORGANISM | Homo sapiens | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| AUTHORS | 1 (bases 1 to 3507) Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,P.S., Wagner,L., Shennan,C.M., Schuler,G.D., Altschul,S.P., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Rha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,L.S., Bosak,S.A., McSwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodrigues,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E., Sutterich,A., Schein,J.E., Jones,S.J., Skalska,U., Marra,M.A. | | | |
| TITLE | Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences | | | |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) | | | |
| PUBMED | 12477932 | | | |
| REFERENCE | 2 (bases 1 to 3507) Strausberg,R. | | | |
| AUTHORS | Direct Submission | | | |
| TITLE | Submitted (08-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | | | |
| JOURNAL | NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nhgri.nih.gov | | | |
| REMARK | Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghghi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Leric,P., McGospi,R., Maduro,Q.L., Masiello,C., Maekeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D. | | | |
| COMMENT | Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 62 Row: a Column: 10 | | | |

| | | | | | |
|----------|--------|---|---------------------|--|---|
| FEATURES | source | 1..3507 | Location/Qualifiers | 1..3507 | This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21955153. |
| | | /organism="Homo sapiens" | | /organism="Homo sapiens" | |
| | | /mol_type="mRNA" | | /mol_type="mRNA" | |
| | | /db_xref="taxon:9606" | | /db_xref="taxon:9606" | |
| | | /clone="MGC:40117 IMAGE:5212737" | | /clone="MGC:40117 IMAGE:5212737" | |
| | | /tissue_type="Blood, adult leukocytes" | | /tissue_type="Blood, adult leukocytes" | |
| | | /clone_lib="NIH MGC_118" | | /clone_lib="NIH MGC_118" | |
| | | /lab_host="DH10B" | | /lab_host="DH10B" | |
| | | /note="vector: pCMV-SPORT6" | | /note="vector: pCMV-SPORT6" | |
| gene | | 1..3507 | | 1..3507 | |
| | | /gene="NALP12" | | /gene="NALP12" | |
| | | /note="synonyms: RNO2, PYPAF7" | | /note="synonyms: RNO2, PYPAF7" | |
| | | /db_xref="LocusID:91662" | | /db_xref="LocusID:91662" | |
| CDS | | 120..3305 | | 120..3305 | |
| | | /gene="NALP12" | | /gene="NALP12" | |
| | | /codon_start=1 | | /codon_start=1 | |
| | | /product="PYRIN-containing APAF1-like protein 7" | | /product="PYRIN-containing APAF1-like protein 7" | |
| | | /protein_id="AAH28069.1" | | /protein_id="AAH28069.1" | |
| | | /db_xref="GI:20380400" | | /db_xref="GI:20380400" | |
| | | /db_xref="LocusID:91662" | | /db_xref="LocusID:91662" | |
| | | /translation="MLRTAGRDGLCRSLTYLELEAVELEKFKKLYLTATLGGKIPWGSNEKAGLEMAQLLTHFGPEANRLALSTFIRINRDLWEGQEDLVRDTPPGPSSIGNOSTCLLEVSIVTPRKDQETVYVRKFRLEMDNRNARLGEVNLVSHVYTRLVKEHSNPMQOQLDTCRKHARTVHOASPIKIEITLPEDEERPEPPRTVMQGAAGIGKSLAHKMLMDWADGLKFOGRFDYLFVINCENMQSATCSCMODLIFSCWPEPSAPQLRLRVERLLFIIDGDELKPSFHDQGWCLCWEKRPTELLINSLIRKLPLAPQLRLRVERLLFIIDGDELKPSFHDQGWCL | | | |

Qy 460 CGGGACACCGGAGGACCGTGGGACACCGAGCTAGCCCATCAAGATAGAGACCCCTCTTT 519
Db 717 CGGGACACCGGAGGACCGTGGGACACCGAGCTAGCCCATCAAGATAGAGACCCCTCTTT 776
Qy 520 GAGCCAGACGAGGAGCGCCCGGAGCCACCGCCGACCGTGGTCAATGCAAGCGCGGAGGG 579
Db 777 GAGCCAGACGAGGAGCGCCCGGAGCCACCGCCGACCGTGGTCAATGCAAGCGCGGAGGG 836
Qy 580 ATAGGCAAGTCCATGCTGGGACACCAAGGTGATGCTGGACTGGCGGACGCGGAGCTCTTC 639
Db 837 ATAGGCAAGTCCATGCTGGGACACCAAGGTGATGCTGGACTGGCGGACGCGGAGCTCTTC 896
Qy 640 CAAGCCAGATTTGATTAATCTCTTCAATCAATCAATGCAAGGAGATGAACCAAGTGCACG 699
Db 897 CAAGCCAGATTTGATTAATCTCTTCAATCAATCAATGCAAGGAGATGAACCAAGTGCACG 956
Qy 700 GAATGCAAGTCCATGCTGGGACACCAAGGTGATGCTGGACTGGCGGACGCGGAGCTCTTC 759
Db 957 GAATGCAAGTCCATGCTGGGACACCAAGGTGATGCTGGACTGGCGGACGCGGAGCTCTTC 1016
Qy 760 GAGCTCATCCGAGTTCCCGAGCGCTCTCTTTTCAATCAATCAATGCAAGGAGATGAACCAAG 819
Db 1017 GAGCTCATCCGAGTTCCCGAGCGCTCTCTTTTCAATCAATCAATGCAAGGAGATGAACCAAG 1076
Qy 820 CTTCTTTTCCAGATCTCTCAGGACCCCTGGTCCCTCTGCTGGGAGGAGAACCGGCCAAG 879
Db 1077 CTTCTTTTCCAGATCTCTCAGGACCCCTGGTCCCTCTGCTGGGAGGAGAACCGGCCAAG 1136
Qy 880 GAGCTGCTCTTAACAGCTTAATTCGGAAGAGTGTCTCTGAGCTATCTTTGCTCATC 939
Db 1137 GAGCTGCTCTTAACAGCTTAATTCGGAAGAGTGTCTCTGAGCTATCTTTGCTCATC 1196
Qy 940 ACCACACGCGCCACCGCTTTGGAGAGCTCCACGCTGTCTGGAGGACCCAGGACATGT 999
Db 1197 ACCACACGCGCCACCGCTTTGGAGAGCTCCACGCTGTCTGGAGGACCCAGGACATGT 1256
Qy 1000 GAGATCTGGGCTCTCTGAGCAGAAAGGAGAACTCTTACAAAGTATTTTCCACAAT 1059
Db 1257 GAGATCTGGGCTCTCTGAGCAGAAAGGAGAACTCTTACAAAGTATTTTCCACAAT 1316
Qy 1060 GCAGAGCAGCGCGGCGCAAGTCTTCAATTAAGTGGAGGACAAAGAGCTCTTCCACCATG 1119
Db 1317 GCAGAGCAGCGCGGCGCAAGTCTTCAATTAAGTGGAGGACAAAGAGCTCTTCCACCATG 1376
Qy 1120 TGTCTGCTCCCTGCTGTGCTGGTGTGTGTAAGTCTGCTCCAGCAGCAGCTGGAGGT 1179
Db 1377 TGTCTGCTCCCTGCTGTGCTGGTGTGTGTAAGTCTGCTCCAGCAGCAGCTGGAGGT 1436
Qy 1180 GGGGGCTGTGTGAGACAGAGCTCCAGGACCACTGCAAGTGTATCATGCTTACCTGCTG 1239
Db 1437 GGGGGCTGTGTGAGACAGAGCTCCAGGACCACTGCAAGTGTATCATGCTTACCTGCTG 1496
Qy 1240 AGTCTGATGCAACCCAGCGCGGGGCGCGCGCTCCAGCCCGCCACCCACCCAGAGAGGG 1299
Db 1497 AGTCTGATGCAACCCAGCGCGGGGCGCGCGCTCCAGCCCGCCACCCACCCAGAGAGGG 1556
Qy 1300 TTGTGCTCTTGGCGGAGAGTGGCTCTGGAATCAGAAATCTTATTTGAGGAGCAGGAC 1359
Db 1557 TTGTGCTCTTGGCGGAGAGTGGCTCTGGAATCAGAAATCTTATTTGAGGAGCAGGAC 1616
Qy 1360 CTCCGGAAGCAGCGGCTAGACGGGGAAGAGCTCTGCTCTTCTCTCAACATGAACATCTTC 1419
Db 1617 CTCCGGAAGCAGCGGCTAGACGGGGAAGAGCTCTGCTCTTCTCTCAACATGAACATCTTC 1676
Qy 1420 CAGAGGACATCAACTGTGAGAGTACTACAGCTTCACTGAGTTTCCAGGAATTC 1479
Db 1677 CAGAGGACATCAACTGTGAGAGTACTACAGCTTCACTGAGTTTCCAGGAATTC 1736
Qy 1480 TTTGACGCTATGATATATCTTGGACGAGGGGAGGGCGGGCCAGCCAGACAGGAC 1539
Db 1737 TTTGACGCTATGATATATCTTGGACGAGGGGAGGGCGGGCCAGCCAGACAGGAC 1796
Qy 1540 GTGACACGCGCTGTGACCGGAGTACCGGTTTTCTGAAAGGAGCTTCTCTGGCACTCACACG 1599
Db 1797 GTGACACGCGCTGTGACCGGAGTACCGGTTTTCTGAAAGGAGCTTCTCTGGCACTCACACG 1856
Qy 1600 CGCTCTCTGTTTGAATCTCTGAAAGGAGCTTCTCTGGCACTCACACG 1659
Db 1857 CGCTCTCTGTTTGAATCTCTGAAAGGAGCTTCTCTGGCACTCACACG 1916
Qy 1660 TGGAGGTCTCGCGGACATCAAGATGAGCTGTTGAGTGGATCCAAAGCAAGTCTCAG 1719
Db 1917 TGGAGGTCTCGCGGACATCAAGATGAGCTGTTGAGTGGATCCAAAGCAAGTCTCAG 1976
Qy 1720 AGCGACGCGCTCCACCTCTGACAGCGGCTCTTGAAGTCTTTCAGCTGCTTGTACGAGATC 1779
Db 1977 AGCGACGCGCTCCACCTCTGACAGCGGCTCTTGAAGTCTTTCAGCTGCTTGTACGAGATC 2036
Qy 1780 CAGGAGGAGGATTTATCCAGCAGCGCTCTGAGCCATCTTCAGGTGATCTGGTCTCAGCAAC 1839
Db 2037 CAGGAGGAGGATTTATCCAGCAGCGCTCTGAGCCATCTTCAGGTGATCTGGTCTCAGCAAC 2096
Qy 1840 ATTGCTCTCAAGATGAGACACATGCTCTCTGTTCTGTGAAGCGCTGAGAGCGCC 1899
Db 2097 ATTGCTCTCAAGATGAGACACATGCTCTCTGTTCTGTGAAGCGCTGAGAGCGCC 2156
Qy 1900 CAGGTGCTGCACTTGTATGCGGCCACCTACAGCGCGGACCGGAGAGCCGCGCAGGTC 1959
Db 2157 CAGGTGCTGCACTTGTATGCGGCCACCTACAGCGCGGACCGGAGAGCCGCGCAGGTC 2216
Qy 1960 TCCGACAGAGGCGCACACCTGTTGTCAGCTCAGACCAAGAGAGACCGTCTCTCTGAGAC 2019
Db 2217 TCCGACAGAGGCGCACACCTGTTGTCAGCTCAGACCAAGAGAGACCGTCTCTCTGAGAC 2273
Qy 2020 GCTTACAGTGAAATCTGCGCAGCGCCCTGTGCAACCAATCCAACTGATAGAGCTGTCT 2079
Db 2274 GCTTACAGTGAAATCTGCGCAGCGCCCTGTGCAACCAATCCAACTGATAGAGCTGTCT 2333
Qy 2080 CTGTACCGAATGCGCTGCGCAGCGCGGGGTGAAGCTGCTGTCTCAAGAGCTCAGACAC 2139
Db 2334 CTGTACCGAATGCGCTGCGCAGCGCGGGGTGAAGCTGCTGTCTCAAGAGCTCAGACAC 2393
Qy 2140 CCAATCTGCAAACTTTCAGAACTGAGGCTGAAGAGGTCGCGCATCTCAGCTCAGCCTGC 2199
Db 2394 CCAATCTGCAAACTTTCAGAACTGAGGCTGAAGAGGTCGCGCATCTCAGCTCAGCCTGC 2453
Qy 2200 GAGGACCTCTCTGACGCTCTCATAGCAATGAATTTGACAAGATGATCTCAGTGGC 2259
Db 2454 GAGGACCTCTCTGACGCTCTCATAGCAATGAATTTGACAAGATGATCTCAGTGGC 2513
Qy 2260 AACGGCTTGGATTTCCAGGCGCATGCTGCTTTGCGAGGGCTGCGGCATCCCAATGC 2319
Db 2514 AACGGCTTGGATTTCCAGGCGCATGCTGCTTTGCGAGGGCTGCGGCATCCCAATGC 2573
Qy 2320 AGGCTGCGAGATGATTCAGTTGAGGAAATGTCAGCTGAGTCCGGGCTTGTGAGGAGATG 2379
Db 2574 AGGCTGCGAGATGATTCAGTTGAGGAAATGTCAGCTGAGTCCGGGCTTGTGAGGAGATG 2633
Qy 2380 GCTTCTGCTGCTCGGACCAACCCACATCTGTTGAGTTGGAACCTGACAGAAATGCACTG 2439
Db 2634 GCTTCTGCTGCTCGGACCAACCCACATCTGTTGAGTTGGAACCTGACAGAAATGCACTG 2693
Qy 2440 GAGGATTTGGGCTGAGGTTTATATGCGAGGGAATGAGGACCCAGCTGTCAGACTACCG 2499
Db 2694 GAGGATTTGGGCTGAGGTTTATATGCGAGGGAATGAGGACCCAGCTGTCAGACTACCG 2753
Qy 2500 ACTTTGCTGCTGAGATCTGCGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2559
Db 2754 ACTTTGCTGCTGAGATCTGCGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2813
Qy 2560 CTGAGTGTGAACAGAGCTGAGAGAGCTGGAACCTGAGCCCTGAGCTGAGTGTGGGGGACCTC 2619
Db 2814 CTGAGTGTGAACAGAGCTGAGAGAGCTGGAACCTGAGCCCTGAGTGTGGGGGACCTC 2873
Qy 2620 GGGTGTGCTGCTGCTGAGGGGCTCAGGCTCCAGCTGCAAGCTTCCAGACCTCTGCGG 2679


```
Db 2874 GGGGTGCTGCTGCTGTGTGAGGGCTCAGGCATCCACGTGCAAGCTCCAGACCTCCAGACCTGGCG 2933
Qy 2680 TTGGGCATCTGCGGCTGGGCTCTGCGGCTGTGAGGGTCTTTCTGTGGTCTCCAGGCC 2739
Db 2934 TTGGGCATCTGCGGCTGGGCTCTGCGGCTGTGAGGGTCTTTCTGTGGTCTCCAGGCC 2993
Qy 2740 AACCAACACTCCGGGAGCTGACCTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGG 2799
Db 2994 AACCAACACTCCGGGAGCTGACCTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGG 3053
Qy 2800 TTGCTGGCTGAGGGCTGCAACATCCCGCTGCAGCTCCAGAACTGCTGGCTGGATAGC 2859
Db 3054 TTGCTGGCTGAGGGCTGCAACATCCCGCTGCAGCTCCAGAACTGCTGGCTGGATAGC 3113
Qy 2860 TTGCTGGCTGAGGGCTGCAACATCCCGCTGCAGCTCCAGAACTGCTGGCTGGATAGC 2919
Db 3114 TTGCTGGCTGAGGGCTGCAACATCCCGCTGCAGCTCCAGAACTGCTGGCTGGATAGC 3173
Qy 2920 TTGACCGACCTTACCTGACCAACACCGCTTAGGGGACACAGGTGCTGGCTTGGC 2979
Db 3174 TTGACCGACCTTACCTGACCAACACCGCTTAGGGGACACAGGTGCTGGCTTGGC 3233
Qy 2980 AAGCGGCTGAGCCATCTCGGCTGCAAACTCCGAGTCCCTGTTATTTGGGATGACCTG 3039
Db 3234 AAGCGGCTGAGCCATCTCGGCTGCAAACTCCGAGTCCCTGTTATTTGGGATGACCTG 3293
Qy 3040 AATAAAATGACCCACAGTAGGTGGCAGCGCTTCGAGTAAACAAACCTTATTGGACATT 3099
Db 3294 AATAAAATGACCCACAGTAGGTGGCAGCGCTTCGAGTAAACAAACCTTATTGGACATT 3353
Qy 3100 GGCTGCTGA 3108
Db 3354 GGCTGCTGA 3362

RESULT 6
AY154467 3827 bp mRNA linear PRI 20-FEB-2003
LOCUS Homo sapiens NALP12 (NALP12) mRNA, complete cds.
DEFINITION
ACCESSION AY154467
VERSION AY154467.1 GI:28436377
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Tschoep, J., Martinon, F. and Burns, K.
TITLE NALPs: a novel protein family involved in inflammation
JOURNAL Nat. Rev. Mol. Cell Biol. 4 (2), 95-104 (2003)
MEDLINE 22451042
PUBMED 12563287
REFERENCE 2 (bases 1 to 3827)
AUTHORS Martinon, F., Hofmann, K. and Tschopp, J.
TITLE Direct Submision
JOURNAL Submitted (25-SEP-2002) Institute of Biochemistry, University of
Lausanne, ch. des Boveresses 155, Epalinges, VD 1066, Switzerland
FEATURES
source
1..3827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..3827
/gene="NALP12"
177..3362
/gene="NALP12"
/note="member of the NALP protein family involved in
inflammation; contains NACHT, LRR and PYD domains"
/codon_start=1
/product="NALP12"
/protein_id="RAO18163.1"
/db_xref="GI:28436378"
/translation="MLRTAGRDGLRLSLTYLSELEAVELEKFKLYLGTATBELGSKIP
```

ORIGIN

```
Query Match 96.5%; Score 2997.8; DB 9; Length 3827;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 3103; Conservative 0; Mismatches 2; Indels 84; Gaps 2;

Qy 1 ATGCTACGAACCGCAGGAGGACGGCTCTGTGCGCTGTCCACTCTTCTGGAAGAACTC 60
Db 177 ATGCTACGAACCGCAGGAGGACGGCTCTGTGCGCTGTCCACTCTTCTGGAAGAACTC 236
Qy 61 GAGGCTGTGGAACCTGAAGAGTTCAAGTTATACCTTGGGGACCGGACAGAGCTGGGAA 120
Db 237 GAGGCTGTGGAACCTGAAGAGTTCAAGTTATACCTTGGGGACCGGACAGAGCTGGGAA 296
Qy 121 GGCAAGATCCCTTGGGGAGCATGGAGAAAGCCCGTCCCTTGGAAATGGCCAGCTGCTC 180
Db 297 GGCAAGATCCCTTGGGGAGCATGGAGAAAGCCCGTCCCTTGGAAATGGCCAGCTGCTC 356
Qy 181 ATCAACCACTTGGGGACAGAGGAGCTGTGAGGTGGCTCTCAGACCTTTGAGCGGATA 240
Db 357 ATCAACCACTTGGGGACAGAGGAGCTGTGAGGTGGCTCTCAGACCTTTGAGCGGATA 416
Qy 241 AACAGGAAGACCTGTGGGAGAGGACAGAGAGGACCTGTGTGAGG----- 288
Db 417 AACAGGAAGACCTGTGGGAGAGGACAGAGAGGACCTGTGTGAGGAGATACCCCACT 476
Qy 289 ----- 288
Db 477 GGTGGCCGCTCCTCACTTGGGAACCACTCAACATGCCTTCTGGAAGTCTCTCTTGTCACT 536
Qy 289 -----GATCCCAAGGAAACCTA CAGGAGCTATGTCCGAGGAAATTCGGGCTCATG 339
Db 537 CCAAGAAAAGATCCCAAGGAAACCTA CAGGAGCTATGTCCGAGGAAATTCGGGCTCATG 596
Qy 340 GAAGACCGCAATGCGGCTAGGGGAATGTGCAACCTCAGCAGCAGCAGCAGCAGCAGCAGC 399
Db 597 GAAGACCGCAATGCGGCTAGGGGAATGTGCAACCTCAGCAGCAGCAGCAGCAGCAGCAGC 656
Qy 400 CTGCTGTGTGAAGAGCACTCAAAACCCATG CAGGTCCAGCAGCAGCTTCTGAGACACAGGC 459
Db 657 CTGCTGTGTGAAGAGCACTCAAAACCCATG CAGGTCCAGCAGCAGCTTCTGAGACACAGGC 716
Qy 460 CGGGGACACGCGAGGACCGTGGGACAC CAGGCTAGCCCCATCAAGATAGAGACCTCTTT 519
Db 717 CGGGGACACGCGAGGACCGTGGGACAC CAGGCTAGCCCCATCAAGATAGAGACCTCTTT 776
Qy 520 GAGCCAGACGAGGAGCGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 579
Db 777 GAGCCAGACGAGGAGCGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 836
Qy 580 ATAGGCAAGTCCATGCTGGCACACAAGGTGATGTGACTGGCGGACGCGGAAGCTCTTC 639
Db 837 ATAGGCAAGTCCATGCTGGCACACAAGGTGATGTGACTGGCGGACGCGGAAGCTCTTC 896
Qy 640 CAAGGCAGATTGATTATCTCTTCTCATCACTCAACTGCAGGGAGATGAACACAGAGTGCCACG 699
```


| | | | |
|------|----|--|------|
| 1061 | Db | GAGCTCATCCGAGTTTCCCGAGCGGCTCTCTTTTATCATCATGACGGCTTTCGATGAGCTCAAG | 1120 |
| 820 | Qy | CCTTTCTTTCCAGATCTCTCAGGGACCCCTGGTGCCTCTGCTGGGAGGAGAAACCGGCCACG | 879 |
| 1121 | Db | CCTTTCTTTCCAGATCTCTCAGGGACCCCTGGTGCCTCTGCTGGGAGGAGAAACCGGCCACG | 1180 |
| 880 | Qy | GAGCTGCTTTTAAACAGCTTAAATTCGGGAAGAGCTGCTCCTTGAGCTATCTTTGCTCATC | 939 |
| 1181 | Db | GAGCTGCTTTTAAACAGCTTAAATTCGGGAAGAGCTGCTCCTTGAGCTATCTTTGCTCATC | 1240 |
| 940 | Qy | ACACAGCGGCCACCGCTTTGGAGAGCTCCACCGTCTGCTGGAGACCCACAGCATGTG | 999 |
| 1241 | Db | ACCACAGCGGCCACCGCTTTGGAGAGCTCCACCGTCTGCTGGAGACCCACAGCATGTG | 1300 |
| 1000 | Qy | GAGATCCTGGCTTCTCTGAGCGAAGAGGAGGAATACTTCTCAAGATATTTCCAAAT | 1059 |
| 1301 | Db | GAGATCCTGGCTTCTCTGAGCGAAGAGGAGGAATACTTCTCAAGATATTTCCAAAT | 1360 |
| 1060 | Qy | GCAGACAGCGGGCCAGTCTTCAATTAAGTGGAGCAACAGAGCTCTCTTCAACATG | 1119 |
| 1361 | Db | GCAGACAGCGGGCCAGTCTTCAATTAAGTGGAGCAACAGAGCTCTCTTCAACATG | 1420 |
| 1120 | Qy | TGCTTGCCTCCCTGGTGTGCTGGGTGTGTACTCTCCCTCCAGACAGCTGGAGGT | 1179 |
| 1421 | Db | TGCTTGCCTCCCTGGTGTGCTGGGTGTGTACTCTCCCTCCAGACAGCTGGAGGT | 1480 |
| 1180 | Qy | GGGGGGCTGTTGAGACAGAGCTCCAGGACCACTGCACTGAGTGTATGCTCTACTCTG | 1239 |
| 1481 | Db | GGGGGGCTGTTGAGACAGAGCTCCAGGACCACTGCACTGAGTGTATGCTCTACTCTG | 1540 |
| 1240 | Qy | AGTCTGATCAACCCAGCGGGGGCCCGCGCTCCAGCCCCCAACCCAGAGAGG | 1299 |
| 1541 | Db | AGTCTGATCAACCCAGCGGGGGCCCGCGCTCCAGCCCCCAACCCAGAGAGG | 1600 |
| 1300 | Qy | TTGGTCTCTTGGCGGAGATGGGCTCTGGATCAGAAATCTTATTTGAGGAGCAGGAC | 1359 |
| 1601 | Db | TTGGTCTCTTGGCGGAGATGGGCTCTGGATCAGAAATCTTATTTGAGGAGCAGGAC | 1660 |
| 1360 | Qy | CTCCGGAAGCACCGCTTAGACGGGGAAGACGCTCTGCTTCTCTCAATGAAATCTTC | 1419 |
| 1661 | Db | CTCCGGAAGCACCGCTTAGACGGGGAAGACGCTCTGCTTCTCTCAATGAAATCTTC | 1720 |
| 1420 | Qy | CAGAAGACATCAACTGTGAGAGTACTACAGCTTCATCCACTGAGTTTCCAGGAATC | 1479 |
| 1721 | Db | CAGAAGACATCAACTGTGAGAGTACTACAGCTTCATCCACTGAGTTTCCAGGAATC | 1780 |
| 1480 | Qy | TTTCAGCTATGACTATATCTCTGACGAGGGGGGGGGGAGGCCAGGCCAGACGAGC | 1539 |
| 1781 | Db | TTTCAGCTATGACTATATCTCTGACGAGGGGGGGGGGAGGCCAGGCCAGACGAGC | 1840 |
| 1540 | Qy | GTGACCAAGCTGTTGACCGAGTACGGGTTTTCTGAAAGAGGCTTCTGGCATCAACGAC | 1599 |
| 1841 | Db | GTGACCAAGCTGTTGACCGAGTACGGGTTTTCTGAAAGAGGCTTCTGGCATCAACGAC | 1900 |
| 1600 | Qy | CGCTTCTGTTTGAATCTCTGAAACGAGGACACGAGGCCACTGAGGAAGAGTCTCTGC | 1659 |
| 1901 | Db | CGCTTCTGTTTGAATCTCTGAAACGAGGACACGAGGCCACTGAGGAAGAGTCTCTGC | 1960 |
| 1660 | Qy | TGGAGAGTCTCGCGCACATCAAGATGGAACCTGTTGAGTGGATCCAAAGCAAAAGCTCAG | 1719 |
| 1961 | Db | TGGAGAGTCTCGCGCACATCAAGATGGAACCTGTTGAGTGGATCCAAAGCAAAAGCTCAG | 2020 |
| 1720 | Qy | AGCGACGCTCCACCTGACAGAGGCTCTTGGAGTCTTACGCTGCTTGTACGAGATC | 1779 |
| 2021 | Db | AGCGACGCTCCACCTGACAGAGGCTCTTGGAGTCTTACGCTGCTTGTACGAGATC | 2080 |
| 1780 | Qy | CAGGAGGAGGATTTATCCAGACGCTCTGAGCCACTTCCAGCTGATCGTGGTCAACAC | 1839 |
| 2081 | Db | CAGGAGGAGGATTTATCCAGACGCTCTGAGCCACTTCCAGCTGATCGTGGTCAACAC | 2140 |
| 1840 | Qy | ATTGCTCTCAAGATGGAGCACATGGTCTCTCTGCTCTGATGAGCGCTGACGAGCGCC | 1899 |

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and Maseho, Y.

AUTHORS

Full-length cDNA sequences

Patent: BP 1347046-A 718 24-SEP-2003;

Research Association for Biotechnology (JP)

JOURNAL

Location/Qualifiers

1. 3466

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

FEATURES

source

Query Match 85.1%; Score 2644.2; DB 6; Length 3466;

Best Local Similarity 91.9%; Pred. No. 0;

Matches 2931; Conservative 0; Mismatches 3; Indels 255; Gaps 3;

Qy 1 ATGCTACGAACCGCAGGAGGAGCGGCTCTGTGCTGCTCCACTACTTGGAAAGACTC 60

Db 145 ATGCTACGAACCGCAGGAGGAGCGGCTCTGTGCTGCTCCACTACTTGGAAAGACTC 204

Qy 61 GAGCTGTGGAACCTGAAGAGTTCAAGTTATACCTGGGACCGCAGACAGCTGGAGAA 120

Db 205 GAGCTGTGGAACCTGAAGAGTTCAAGTTATACCTGGGACCGCAGACAGCTGGAGAA 264

Qy 121 GCGAAGATCCCTCGGGAAGCATGGAGAGGCGGCTCCCTGGAAATGGCCACGCTGCTC 180

Db 265 GCGAAGATCCCTCGGGAAGCATGGAGAGGCGGCTCCCTGGAAATGGCCACGCTGCTC 324

Qy 181 ATCACCACTTCGGGCCAGAGGAGGCGCTGGAGTTGGCTCTCAGCACTTTGACGGATA 240

Db 325 ATCACCACTTCGGGCCAGAGGAGGCGCTGGAGTTGGCTCTCAGCACTTTGACGGATA 384

Qy 241 AACAGGAGGACCTGTGGGAGAGGACAGAGAGGACCTGGTGAGG----- 288

Db 385 AACAGGAGGACCTGTGGGAGAGGACAGAGAGGACCTGGTGAGGAGTACCCACCT 444

Qy 289 ----- 288

Db 445 GGTGGCCGCTCCTCACTTGGGAACCACTCAACATGCCTTCTGGAAGTCTCTTGTCACT 504

Qy 289 -----GATCCCCAGGAACCTACAGGAGCTATGTCCGAGGAAATTCGGCTCATG 339

Db 505 CCAAGAAAAGATCCCAGGAACCTACAGGAGCTATGTCCGAGGAAATTCGGCTCATG 564

Qy 340 GAACACCGCAATGCGGCTAGGAGGATGTGTAACCTCAGCCACCGGTACACCGGCTC 399

Db 565 GAACACCGCAATGCGGCTAGGAGGATGTGTAACCTCAGCCACCGGTACACCGGCTC 624

Qy 400 CTGCTGGTGAAGGAGCACTCAAAACCCATGAGGTCCAGCAGCAGCTTCTGGACACAGGC 459

Db 625 CTGCTGGTGAAGGAGCACTCAAAACCCATGAGGTCCAGCAGCAGCTTCTGGACACAGGC 684

Qy 460 CGGGGACACGCGAGGACCGTGGGACACCAAGGCTAGCCCCCATCAAGATAGAGACCTCTTT 519

Db 685 CGGGGACACGCGAGGACCGTGGGACACCAAGGCTAGCCCCCATCAAGATAGAGACCTCTTT 744

Qy 520 GAGCCAGACGAGGAGCGCCCGAGGACACCGGCGCACCGTGTGTCATGCAAGGCGGCGAGGG 579

Db 745 GAGCCAGACGAGGAGCGCCCGAGGACACCGGCGCACCGTGTGTCATGCAAGGCGGCGAGGG 804

Qy 580 ATAGGCAAGTCCATGCTGGCACACAAGGTATGCTGGACTGGGCGGACGGGAAGCTCTTC 639

Db 805 ATAGGCAAGTCCATGCTGGCACACAAGGTATGCTGGACTGGGCGGACGGGAAGCTCTTC 864

Qy 640 CAAGGCAGATTTGATTATCTCTTCTATCACTCAACTGCAGGGAGATGAACCAAGTGCACG 699

865 CAAGGCAGATTTGATTATCTCTTCTATCACTCAACTGCAGGGAGATGAACACAGAGTGCACG 924

700 GAATGCAGCATGCAAGACCTCATCTTCAGCTGTGGCTGAGCCAGCGGCTCTCCAG 759

925 GAATGCAGCATGCAAGACCTCATCTTCAGCTGTGGCTGAGCCAGCGGCTCTCCAG 984

760 GAGCTCATCCGAGTTCGCCGAGCGCTCTTTTTCATCATCGACGGCTTCGATGAGCTCAAG 819

985 GAGCTCATCCGAGTTCGCCGAGCGCTCTTTTTCATCATCGACGGCTTCGATGAGCTCAAG 1044

820 CTTTCTTTCACAGATCTCTAGGACCTCTGGTCCCTCTGCTGGGAGGAGAAACGGCCACG 879

1045 CTTTCTTTCACAGATCTCTAGGACCTCTGGTCCCTCTGCTGGGAGGAGAAACGGCCACG 1104

880 GAGCTGCTCTTAACAGCTTAATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATC 939

1105 GAGCTGCTCTTAACAGCTTAATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATC 1164

940 ACCACAGCGCCACGGCTTTGGAGAACTCCACCGTCTGCTGGAGCACCCACGAGCATGTG 999

1165 ACCACAGCGCCACGGCTTTGGAGAACTCCACCGTCTGCTGGAGCACCCACGAGCATGTG 1224

1000 GAGATCTCGGCTTCTCTGAGGACAGAAAGGAAGATCTCTTCAAGTATTTCCACAAAT 1059

1225 GAGATCTCGGCTTCTCTGAGGACAGAAAGGAAGATCTCTTCAAGTATTTCCACAAAT 1284

1060 GCAGAGCAGCGGGCCCAAGTCTTCAATTAGCTGAGGACAAACGAGCTCTCTTCCACCAAT 1119

1285 GCAGAGCAGCGGGCCCAAGTCTTCAATTAGCTGAGGACAAACGAGCTCTCTTCCACCAAT 1344

1120 TGCTTCTCCCTCGTGTGCTGCTGGTGTGTGTACCTGCTCCAGCAGCAGCTGAGAGGT 1179

1345 TGCTTCTCCCTCGTGTGCTGCTGGTGTGTGTGTACCTGCTCCAGCAGCAGCTGAGAGGT 1404

1180 GGGGGGCTGTTGAGACAGACAGCTCCAGGACCACTGAGTGATCATGCTCTACTGCTG 1239

1405 GGGGGGCTGTTGAGACAGACAGCTCCAGGACCACTGAGTGATCATGCTCTACTGCTG 1464

1240 AGTCTGATGCAACCCAGCGGGGGGGGGCTCCAGGCTCCAGCCCAACCCACGAGAGGG 1299

1465 AGTCTGATGCAACCCAGCGGGGGGGGGCTCCAGGCTCCAGCCCAACCCACGAGAGGG 1524

1300 TTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAATCTCTATTGAGGAGCAGAGAC 1359

1525 TTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAATCTCTATTGAGGAGCAGAGAC 1584

1360 CTCGGGAAGCA CGGCTTAGA CGGGAAGAGAGCTCTCTGCTCTCTCAACATGAA CATCTTC 1419

1585 CTCGGGAAGCA CGGCTTAGA CGGGAAGAGAGCTCTCTGCTCTCTCAACATGAA CATCTTC 1644

1420 CAGNAGACATCAACTGTGAGAGGTACTACAGCTTTCATCCACTGAGTTCCTCAGGAATTC 1479

1645 CAGNAGACATCAACTGTGAGAGGTACTACAGCTTTCATCCACTGAGTTCCTCAGGAATTC 1704

1480 TTTGCACTATGACTATATATCTCTGGACGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1539

1705 TTTGCACTATGACTATATATCTCTGGACGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1764

1540 GTGACCAAGGCTGTGACCGAGTACCGCTTTTCTGAAAGGAGCTTCCTGGCACTCACCAGC 1599

1765 GTGACCAAGGCTGTGACCGAGTACCGCTTTTCTGAAAGGAGCTTCCTGGCACTCACCAGC 1824

1600 CGCTTCTGCTTGGACTCTTGAACGAGGAGCCAGGAGCCACTGGAGAGAGTCTCTGCG 1659

1825 CGCTTCTGCTTGGACTCTTGAACGAGGAGCCAGGAGCCACTGGAGAGAGTCTCTGCG 1884

1660 TGAAGGCTCTCGCGGACATCAAGATGGAAGCTGTTGAGTGGATCCAAAGCAAGAGCTCAG 1719

1885 TGAAGGCTCTCGCGGACATCAAGATGGAAGCTGTTGAGTGGATCCAAAGCAAGAGCTCAG 1944

1720 AGCGAGCGCTCCACCGCTGCAGGAGGGCTCTTGGAGTTCCTTCAAGTCTCTGTACGAGATC 1779

Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Kanehori, K., Sekine, M., Kikuchi, H., Murakawa, K., Kawakami, B., Suzuki, Y., Sugano, S., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Negai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 3466)

JOURNAL
REFERENCE

Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7

JOURNAL
TITLE

Karusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3985)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers

1..3466
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="D9OST2002673"
/cell_type="CD34+ Cells"
/clone_lib="DSOST2"
/note="cloning vector: pME18SFL3-mRNA from CD34+ cells after 9-days ODF induction.-primary culture, CD34+ Cells"

ORIGIN

Query Match 85.1%; Score 2644.2; DB 9; Length 3466;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 2931; Conservative 0; Mismatches 3; Indels 255; Gaps 3;
QY 1 ATGCTACGAACCGCAGGAGGAGCGGCTCTGTGCGCTGTCCACCTACTTGGAGAACTC 60
DB 145 ATGCTACGAACCGCAGGAGGAGCGGCTCTGTGCGCTGTCCACCTACTTGGAGAACTC 204
QY 61 GAGGCTGTGGAACTCAAGAGTTCAAGTTATACCTTGGGAGCGGAGACAGAGCTGGAGAA 120
DB 205 GAGGCTGTGGAACTCAAGAGTTCAAGTTATACCTTGGGAGCGGAGACAGAGCTGGAGAA 264
QY 121 GCGAGATCCCTGGGAGAGCATGAGAGGCGCGTCCCTCGGAAATGGCCACCTGCTC 180
DB 265 GCGAGATCCCTGGGAGAGCATGAGAGGCGCGTCCCTCGGAAATGGCCACCTGCTC 324
QY 181 ATCAACCCATTCGGGCCAGAGAGGCGCTGGAGGTTGGCTCTCAGCACCTTTGAGCGGATA 240
DB 325 ATCAACCCATTCGGGCCAGAGAGGCGCTGGAGGTTGGCTCTCAGCACCTTTGAGCGGATA 384
QY 241 AACAGGAAGACCTGTGGAGAGAGGAGCAGAGAGGACCTGGTGAGG-----288
DB 385 AACAGGAAGACCTGTGGAGAGAGGAGCAGAGAGGAGGACCTGGTGAGGAGATACCCACCT 444
QY 289 -----288
DB 445 GGTGGCCCGTCTCACTTTGGAAACAGTCAACATGCGCTTCTGGGAAGTCTCTTTGTCACT 504
QY 289 -----GATCCCCAGGAAACCTACAGGAGCTATGTCCGAGGAATTCGGGCTCATG 339
DB 505 CCAAGAAAGATCCCGAGAAACCTACAGGAGCTATGTCCGAGGAATTCGGGCTCATG 564
QY 340 GAAGACCGCAATGCGCGCTTAGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTC 399
DB 565 GAAGACCGCAATGCGCGCTTAGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTC 624
QY 400 CTGCTGTGAGGAGCACTCAACCCCATGAGCTCCAGCAGCAGCTTCTGGACACAGGC 459
DB 625 CTGCTGTGAGGAGCACTCAACCCCATGAGCTCCAGCAGCAGCTTCTGGACACAGGC 684

QY 460 CGGGGACACGCGAGGACCGTGGGACACCGAGCTAGCCCACTCAAGATAGAGACCCCTCTTT 519
DB 685 CGGGGACACGCGAGGACCGTGGGACACCGAGCTAGCCCACTCAAGATAGAGACCCCTCTTT 744
QY 520 GAGCCAGACGAGGAGCGCCCGGACCGGCGACCGTGTCTATGCAAGGCGCGCAGGG 579
DB 745 GAGCCAGACGAGGAGCGCCCGGACCGGCGACCGTGTCTATGCAAGGCGCGCAGGG 804
QY 580 ATAGGCAAGTTCATGCTGGGACACAAAGTGATGTGACTGGGCGGAGCGGAAGCTCTTC 639
DB 805 ATAGGCAAGTTCATGCTGGGACACAAAGTGATGTGACTGGGCGGAGCGGAAGCTCTTC 864
QY 640 CAAGGCAAGTTCATGATTTCTTCTTACATCAACTGCGAGGAGATGAACACAGAGTGCCACG 699
DB 865 CAAGGCAAGTTCATGATTTCTTCTTACATCAACTGCGAGGAGATGAACACAGAGTGCCACG 924
QY 700 GAATGACGATGCAAGACCTCATCTTCACTGCTGGCTGAGCCAGGCGGCTCTCCAG 759
DB 925 GAATGACGATGCAAGACCTCATCTTCACTGCTGGCTGAGCCAGGCGGCTCTCCAG 984
QY 760 GAGCTCATCGAGTTCGCGAGCGCTCTCTTTTCACTCATCGAGCGCTTCGATGAGCTCAAG 819
DB 985 GAGCTCATCGAGTTCGCGAGCGCTCTCTTTTCACTCATCGAGCGCTTCGATGAGCTCAAG 1044
QY 820 CCTTCTTTCCACGATCCTCAGGAGCCCTGGTGCCCTCTGCTGGGAGGAGAAACGCGCCACG 879
DB 1045 CCTTCTTTCCACGATCCTCAGGAGCCCTGGTGCCCTCTGCTGGGAGGAGAAACGCGCCACG 1104
QY 880 GAGCTGCTCTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATC 939
DB 1105 GAGCTGCTCTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATC 1164
QY 940 ACCACACGCGCCCGGCTTTGGAGAGCTCCACCGTCTGCTGGAGGACCCCGAGCAGCTG 999
DB 1165 ACCACACGCGCCCGGCTTTGGAGAGCTCCACCGTCTGCTGGAGGACCCCGAGGAGTGTG 1224
QY 1000 GAGATCCTGGGCTTCTCTGAGGACGAGAAAGAAAGAAATCTTCTCAAGTATTTCCACAAT 1059
DB 1225 GAGATCCTGGGCTTCTCTGAGGACGAGAAAGAAAGAAATCTTCTCAAGTATTTCCACAAT 1284
QY 1060 GCAGAGACGCGGCGCAAGTCTTCAATTAAGTGAAGGACAAACGAGGCTCTTCCACCATG 1119
DB 1285 GCAGAGACGCGGCGCAAGTCTTCAATTAAGTGAAGGACAAACGAGGCTCTTCCACCATG 1344
QY 1120 TGCTTCGTCCTGCTGGTGCTGGTGCTGTGTACCTGCTCCAGCAGCAGCTGCGAGGCT 1179
DB 1345 TGCTTCGTCCTGCTGGTGCTGGTGCTGTGTACCTGCTCCAGCAGCAGCTGCGAGGCT 1404
QY 1180 GGGGGGCTGTGAGACAGAGCTCCAGGACCAACCACTGACGTGTATCATGCTCTACCTGCTG 1239
DB 1405 GGGGGGCTGTGAGACAGAGCTCCAGGACCAACCACTGACGTGTATCATGCTCTACCTGCTG 1464
QY 1240 AGTCTGATGCAACCCAAAGCGGGGCGCCCGGCGCTCCAGCCCGCCACCCCAACAGAGAGG 1299
DB 1465 AGTCTGATGCAACCCAAAGCGGGGCGCCCGGCGCTCCAGCCCGCCACCCCAACAGAGAGG 1524
QY 1300 TTGTGCTCTTGGCGGAGATGGGCTCTGGAAATCAGAAATCTTATTTGAGGAGCAGGAC 1359
DB 1525 TTGTGCTCTTGGCGGAGATGGGCTCTGGAAATCAGAAATCTTATTTGAGGAGCAGGAC 1584
QY 1360 CTCGGGAGACGCGCTTAGAGGAGAGAGCTCTCTGCTCTCTTCTTCAACATGAACATCTTC 1419
DB 1585 CTCGGGAGACGCGCTTAGAGGAGAGAGCTCTCTGCTCTCTTCTTCAACATGAACATCTTC 1644
QY 1420 CAGAGGACATCAACTGTGAGAGGTACTACAGCTTCTCACTTGAATTTTCAGGAATTC 1479
DB 1645 CAGAGGACATCAACTGTGAGAGGTACTACAGCTTCTCACTTGAATTTTCAGGAATTC 1704
QY 1480 TTTTGACCTATGCTATATCTTGGAGGAGGCGGGGCGGCGGAGCCCGAGCCAGGAC 1539
DB 1705 TTTTGACCTATGCTATATCTTGGAGGAGGCGGGGCGGCGGAGCCCGAGCCAGGAC 1764
QY 1540 GTGACAGGCTGTGACCGAGTACGCGTCTTTCTGAAAGGAGGCTTCTCTGGCACTCACCAGC 1599

| | | | | | | | |
|----|------|---|------|------------|---|---|------------------------|
| Db | 1765 | GTGACAGGCTGTTGACCGAGTACCGGTTTCTGAAAGGAGCTTCTGCGACTCAACGAC | 1824 | Db | 2842 | GGGGTGTGCTGTGTGTGAGGGGCTCAGGCATCCACGTCGAAAGCTCCAGACCCCTGC | 2899 |
| Qy | 1600 | CGCTTCCTGTTGGACTCTGAACGAGGAGACAGAGCCACTGAGAGAGTCTCTGC | 1659 | Qy | 2680 | TTGGGCATCTGCCGGCTGGGCTCTGCCGCTGTGTGAGGGTCTTCTGTGTGCTCCAGGCC | 2739 |
| Db | 1825 | CGCTTCCTGTTGGACTCTGAACGAGGAGACAGAGCCACTGAGAGAGTCTCTGC | 1884 | Db | 2900 | ----- | 2899 |
| Qy | 1660 | TGGAAGTCTCGCGGCACATCAAGATGGACCTGTTCAGTGGATCCAAAGCAAGCTCAG | 1719 | Qy | 2740 | AACCACAACCTCCGGGAGCTGGACTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGG | 2799 |
| Db | 1885 | TGGAAGTCTCGCGGCACATCAAGATGGACCTGTTCAGTGGATCCAAAGCAAGCTCAG | 1944 | Db | 2900 | ----- | 2899 |
| Qy | 1720 | AGCGACGGCTCCACCTGCGAGCGGCTCTTGGAGTTCTTCACTGCTTGTACGAGATC | 1779 | Qy | 2800 | TTGCTGGCTGAGGGCTGCAACATCCCGCTGCAGACTCCAGAACTGTGCTGGATAGC | 2859 |
| Db | 1945 | AGCGACGGCTCCACCTGCGAGCGGCTCTTGGAGTTCTTCACTGCTTGTACGAGATC | 2004 | Db | 2900 | ----- | 2910 |
| Qy | 1780 | CAGAGGAGGATTTATCCAGCAGCGCTGAGCCACTTCCAGCTGATCGTGGTTCAGCAAC | 1839 | Qy | 2860 | TGTGGCTCAGAGCCAGGCTTGTGAGAACTTTACTTTCACCTTGGGGATCAACAGACC | 2919 |
| Db | 2005 | CAGAGGAGGATTTATCCAGCAGCGCTGAGCCACTTCCAGCTGATCGTGGTTCAGCAAC | 2064 | Db | 2911 | TGTGGCTCAGAGCCAGGCTTGTGAGAACTTTACTTTCACCTTGGGGATCAACAGACC | 2970 |
| Qy | 1840 | ATTGCTTCCAAAGATGGAGCAATGGTCTCTCTGTTCTGTCTGAAGCGCTGCAGAGCGCC | 1899 | Qy | 2920 | TTGACCGACCTTTTACCTGACCAACAGCCCTTAGGGGACACAGGTGTCGGACTGCTTGC | 2979 |
| Db | 2065 | ATTGCTTCCAAAGATGGAGCAATGGTCTCTCTGTTCTGTCTGAAGCGCTGCAGAGCGCC | 2124 | Db | 2971 | TTGACCGACCTTTTACCTGACCAACAGCCCTTAGGGGACACAGGTGTCGGACTGCTTGC | 3030 |
| Qy | 1900 | CAGGTGCTGCACTTGTATGGCGCCACTACAGCGCGACGCGGAAGACCGCGAGGTGC | 1959 | Qy | 2980 | AAGCGCTGAGGCATCTCTGGCTGCAAACTCCAGTCTCTCTGTTATTGGGATGAGACCTG | 3039 |
| Db | 2125 | CAGGTGCTGCACTTGTATGGCGCCACTACAGCGCGACGCGGAAGACCGCGAGGTGC | 2184 | Db | 3031 | AAGCGCTGAGGCATCTCTGGCTGCAAACTCCAGTCTCTCTGTTATTGGGATGAGACCTG | 3090 |
| Qy | 1960 | TCCGACGAGCGCACACCTGTTGTGTCAGCTCAGACACAGAGGACCGTTCCTGCTGAC | 2019 | Qy | 3040 | AATAAAATGACCCACAGTAGGTGTCAGCGCTTCAGTAAACAAACCTTATTGGACATT | 3099 |
| Db | 2185 | TCCGACGAGCGCACACCTGTTGTGTCAGCT---ACCAGAGGACCGTTCCTGCTGAC | 2241 | Db | 3091 | AATAAAATGACCCACAGTAGGTGTCAGCGCTTCAGTAAACAAACCTTATTGGACATT | 3150 |
| Qy | 2020 | GCCTACAGTGAACATCTGGAGCGGCTGTGTCACCAATCCAAACCTGTAGAGCTGCT | 2079 | Qy | 3100 | GGCTGCTGA 3108 | |
| Db | 2242 | GCCTACAGTGAACATCTGGAGCGGCTGTGTCACCAATCCAAACCTGTAGAGCTGCT | 2301 | Db | 3151 | GGCTGCTGA 3159 | |
| Qy | 2080 | CTGTACCGAAATGCCCTGGGAGCGGGGTGAAGCTGCTCTGTCAAGGACTCAGACAC | 2139 | RESULT 10 | | | |
| Db | 2302 | CTGTACCGAAATGCCCTGGGAGCGGGGTGAAGCTGCTCTGTCAAGGACTCAGACAC | 2361 | AY116206 | 3395 bp | mRNA | linear PRI 21-MAY-2003 |
| Qy | 2140 | CCCACTGCAAACTTTCAGAACCTGAGGTGAGAGGTGCGGATCTCCAGTCAAGCTGC | 2199 | LOCUS | Homo sapiens monarch-1 splice form III | mRNA, complete cds; | |
| Db | 2362 | CCCACTGCAAACTTTCAGAACCTGAGGTGAGAGGTGCGGATCTCCAGTCAAGCTGC | 2421 | DEFINITION | alternatively spliced. | | |
| Qy | 2200 | GAGGACCTCTGCGAGCTCTATAGCCATAGAAATTTGACAGGATGATCTCAGTGGC | 2259 | ACCESSION | AY116206 | | |
| Db | 2422 | GAGGACCTCTGCGAGCTCTATAGCCATAGAAATTTGACAGGATGATCTCAGTGGC | 2481 | VERSION | AY116206.1 | GI:21711824 | |
| Qy | 2260 | AACGGCTTGGATTCCAGGACATGATGCTGCTTTGCGAGGGCTCCGGCATCCCCAGTGC | 2319 | KEYWORDS | | | |
| Db | 2482 | AACGGCTTGGATTCCAGGACATGATGCTGCTTTGCGAGGGCTCCGGCATCCCCAATGC | 2541 | ORGANISM | Homo sapiens (human) | | |
| Qy | 2320 | AGGCTGCAGATGATTCAGTTGAGGAAGTGTAGCTGAGTCCGGGCTTGTCCAGGAGATG | 2379 | REFERENCE | | | |
| Db | 2542 | AGGCTGCAGATGATTCAGTTGAGGAAGTGTAGCTGAGTCCGGGCTTGTCCAGGAGATG | 2601 | AUTHORS | Williams, K.L., Linhoff, M.W., Reed, W., and Ting, J.P.Y. | | |
| Qy | 2380 | GCTTCTGCTCGGACCAACCAATCTGTTGAGTGGACTGACAGGAATGCACCTG | 2439 | TITLE | Monarch-1: A Pyrin/Nucleotide-Binding Domain/Leucine-Rich Repeat Protein That Controls Classical and Nonclassical MHC Class I Genes | | |
| Db | 2602 | GCTTCTGCTCGGACCAACCAATCTGTTGAGTGGACTGACAGGAATGCACCTG | 2661 | JOURNAL | J. Immunol. 170 (11), 5354-5358 (2003) | | |
| Qy | 2440 | GAGGATTTGGGCTGAGGTTATATGCGAGGAGCTGAGGACCCAGTCTGAGACTACGG | 2499 | PUBMED | 12759408 | | |
| Db | 2662 | GAGGATTTGGGCTGAGGTTATATGCGAGGAGCTGAGGACCCAGTCTGAGACTACGG | 2721 | REFERENCE | 2 (bases 1 to 3395) | | |
| Qy | 2500 | ACTTTGGCTGAGTCTGCGGCTCAGTCTGCTGCTGTGACGAGCTGGCTCAACT | 2559 | AUTHORS | Williams, K.L., Linhoff, M.W. and Ting, J.P.Y. | | |
| Db | 2722 | ACTTTGGCTGAGTCTGCGGCTCAGTCTGCTGCTGTGACGAGCTGGCTCAACT | 2781 | TITLE | Submitted (29-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm Road, Chapel Hill, NC 27599, USA | | |
| Qy | 2560 | CTCAGTGTGAACAGAGCTGAGAGCTGAGACTGAGCTGAGCTGAATGAGCTGGGGACCTC | 2619 | JOURNAL | Location/Qualifiers | | |
| Db | 2782 | CTCAGTGTGAACAGAGCTGAGAGCTGAGACTGAGCTGAGCTGAATGAGCTGGGGACCTC | 2841 | FEATURES | 1. .3395 | | |
| Qy | 2620 | GGGTGCTGCTGTGTGAGGCTCAGGACCTCCAGCTGCAAGCTCCAGACCTTCGCGG | 2679 | source | /organism="Homo sapiens" | | |
| | | | | | /mol_type="mRNA" | | |
| | | | | | /db_xref="taxon:9606" | | |
| | | | | | /chromosome="19" | | |
| | | | | | /map="19q13.4" | | |
| | | | | | 221. .3073 | | |
| | | | | | /note="contains alternatively spliced exons 7 and 8" | | |
| | | | | | /codon_start=1 | | |
| | | | | | /product="monarch-1 splice form III" | | |
| | | | | | /protein_id="AA075144.1" | | |
| | | | | | /db_xref="GI:21711825" | | |
| | | | | | /translation="MLTAGRDGICRLSTYLEAVELEKFKLILGTTATLGLGKIP | | |
| | | | | | WGSMEKAGPLEMAQLLTHFGPEAWRLALSTPERINRKLWERGQEDLVRDTPPG | | |

| | |
|---|---|
| PSSIGNOSTCLLEVLVTPRKDPQETRYDVRKFRFLMRDLNRLBSCVNLSHRYTEL | |
| LLVKEHNSPMVOQOQLDTRGRHARTVHQASPIKIELTFEPDEERPEPRTVVMQGA | |
| AGLKEHNSPMVOQOQLDTRGRHARTVHQASPIKIELTFEPDEERPEPRTVVMQGA | |
| APLQELIRVPRFLFIIDGDELKPSFHDPOGWCLEWEEKRPTELLNLSLRKLLP | |
| ELSLITRPTALEKLRLLEHPRVILGFSSEARKEYEVKYPHNAEQAGOVENVYR | |
| DNEPLTFMCFVPLVWVUCVLCLOQLBGGLLRQTSRTTAYVYMLYLLSLMQPKGAP | |
| RLQFPNQRGLCSLAADGLWNQKLFBEQDURKRGLDGEDVSALFNWNI FOKDINCBR | |
| YYSFIHLSFQBFPAAMYIYLDEGGGAGPDQDVTLLTEYAFSERSFALTSLRFLGL | |
| LNEBTRHLEKSLKWSVPHIKMDLQWIOSKQSDGSLTQQGSLBFFSCLYEIYESE | |
| FIOALSHFOVIVSNASKMHMVSSFLCKRCSAOVLHLYGATYADGEDDRARCSA | |
| GAHTLLVQLRPERVLLDAYSEHLAALCTNPMLIELSLVNALGSRGVKLLCOGLRH | |
| PNCKLQNLRLKRCRISACEDLSAALIANKULTRMDLSNGVGFPGMMLLCEGLRHP | |
| QCRQLMLQRLKRCRISACEDLSAALIANKULTRMDLSNGVGFPGMMLLCEGLRHP | |
| CRLRLMLWLDSCGLTACENLYFTLIGINQTLTDLYLTNNALGDVTGRLLCRLSLHP | |
| GCKURVLWLFMDLNMKTHSRALALRVTKPYLDIGC" | |
| 869..892 | |
| /note="Region: P-loop" | |
| 2069..3070 | |
| /note="Region: leucine rich repeat" | |
| 221..289 | |
| /note="Region: pyrin domain" | |
| ORIGIN | |
| Query Match 79.5%; Score 2470; DB 9; Length 3395; | |
| Best Local Similarity 94.5%; Pred. No. 0; | |
| Matches 2616; Conservative 0; Mismatches 65; Indels 87; Gaps 2; | |
| Qy 1 | ATGCTAGCAACCGCAGGCGCGCTCTGTCGCTGTCCACTTCTTGGGAAGACTC 60 |
| Db | |
| Qy 221 | ATGCTAGCAACCGCAGGCGCGCTCTGTCGCTGTCCACTTCTTGGGAAGACTC 280 |
| Db | |
| Qy 61 | GAGCTGTGGAAGCTTCAAGTTATACCTTGGGGACCGCGACAGAGCTGGGAGAA 120 |
| Db | |
| Qy 281 | GAGCTGTGGAAGCTTCAAGTTATACCTTGGGGACCGCGACAGAGCTGGGAGAA 340 |
| Db | |
| Qy 121 | GGCAAGATCCCTGGGAGAGCATGGAGAGCCCGTCCCTGGAAATGGCCAGCTGCTC 180 |
| Db | |
| Qy 341 | GGCAAGATCCCTGGGAGAGCATGGAGAGCCCGTCCCTGGAAATGGCCAGCTGCTC 400 |
| Db | |
| Qy 181 | ATCACCACTTCGGGACAGAGGCGCTTGGAGTTGGCTCTCAGCACCTTTGAGCGGATA 240 |
| Db | |
| Qy 401 | ATCACCACTTCGGGACAGAGGCGCTTGGAGTTGGCTCTCAGCACCTTTGAGCGGATA 460 |
| Db | |
| Qy 241 | AACAGGAAGGACCTGTGGGAGAGGACAGAGAGGAGGACCTGTGTGAGG- 288 |
| Db | |
| Qy 461 | AACAGGAAGGACCTGTGGGAGAGGACAGAGAGGAGGACCTGTGTGAGGATACCCACCT 520 |
| Db | |
| Qy 289 | ----- 288 |
| Db | |
| Qy 521 | GGTGGCCCGTCTCACTTGGGAACCAAGTCAACATGCGCTTCTGGAAGTCTCTCTTGTCACT 580 |
| Db | |
| Qy 289 | -----GATCCCCAGGAACCTACAGGAGCTATGTCCGACGAAATTCGGGCTCATG 339 |
| Db | |
| Qy 581 | CCAAGAAAGATCCCCAGGAACCTACAGGAGCTATGTCCGACGAAATTCGGGCTCATG 640 |
| Db | |
| Qy 340 | GAAGACCGCAATGCGCGCTAGGGGAATGTGTCAACCTTCAGCCACCGGCTACACCGGCTC 399 |
| Db | |
| Qy 641 | GAAGACCGCAATGCGCGCTAGGGGAATGTGTCAACCTTCAGCCACCGGCTACACCGGCTC 700 |
| Db | |
| Qy 400 | CTGCTGTGAAGAGGACCTCAAAACCCATGAGGTTCAGAGAGGAGTTCGGAACAGAGG 459 |
| Db | |
| Qy 701 | CTGCTGTGAAGAGGACCTCAAAACCCATGAGGTTCAGAGAGGAGTTCGGAACAGAGG 760 |
| Db | |
| Qy 460 | CGGGACACGCGAGGACCGTGGGACACAGGCTAGCCCTCAAGATAGAGACCTCTTT 519 |
| Db | |
| Qy 761 | CGGGACACGCGAGGACCGTGGGACACAGGCTAGCCCTCAAGATAGAGACCTCTTT 820 |
| Db | |
| Qy 520 | GAGCCAGACGAGGAGCGCCCGAGCCACCGGCGACCGTGTGTATGCAAGGCGCGGAGGG 579 |
| Db | |
| Qy 821 | GAGCCAGACGAGGAGCGCCCGAGCCACCGGCGACCGTGTGTATGCAAGGCGCGGAGGG 880 |
| Db | |
| Qy 580 | ATAGGCAAGTCCATGTGGCAACAAGGTGATGTGAGCTGGGCGGACGGGAAGCTCTTC 639 |
| Db | |

| | | |
|----|------|---|
| Db | 881 | ATAGGCAAGTCCATGTCTGGCACACAAGGTGATGTGAGCTGGGCGGACGGGAAGCTCTTC 940 |
| Qy | 640 | CAAGGCAGATTGATTATCTCTTACATCAACTGACGGGAGATGAACACAGAGTGCCACG 699 |
| Db | 941 | CAAGGCAGATTGATTATCTCTTACATCAACTGACGGGAGATGAACACAGAGTGCCACG 1000 |
| Qy | 700 | GAATGACGATGCAAGACCTCATCTTCACTGTCTGGCTGAGCCAGCGGCTCTCCAG 759 |
| Db | 1001 | GAATGACGATGCAAGACCTCATCTTCACTGTCTGGCTGAGCCAGCGGCTCTCCAG 1060 |
| Qy | 760 | GAGCTCATCGAGTTCCGAGCGCTCTTCACTGTCTGGCTGAGCCAGCGGCTCTCCAG 819 |
| Db | 1061 | GAGCTCATCGAGTTCCGAGCGCTCTTCACTGTCTGGCTGAGCCAGCGGCTCTCCAG 1120 |
| Qy | 820 | CTTCTTTCACGATCTCAGGAGACCTGTGTCTCTGTCTGGGAGGAGAAACGGCCACG 879 |
| Db | 1121 | CTTCTTTCACGATCTCAGGAGACCTGTGTCTCTGTCTGGGAGGAGAAACGGCCACG 1180 |
| Qy | 880 | GAGCTGCTTCTTAACAGCTTAATTCGGAAGAGCTGTCTCTGTAGCTATCTTTGCTCATC 939 |
| Db | 1181 | GAGCTGCTTCTTAACAGCTTAATTCGGAAGAGCTGTCTCTGTAGCTATCTTTGCTCATC 1240 |
| Qy | 940 | ACCACACGCGCCACGGCTTGGAGAGCTCCACCGTCTGTCTGGAGCACCCCGGACGTG 999 |
| Db | 1241 | ACCACACGCGCCACGGCTTGGAGAGCTCCACCGTCTGTCTGGAGCACCCCGGACGTG 1300 |
| Qy | 1000 | GAGATCTGGGCTTCTCTGAGGACGAAAGGAAGAACTTCTTCAAGTATTTCCACAA 1059 |
| Db | 1301 | GAGATCTGGGCTTCTCTGAGGACGAAAGGAAGAACTTCTTCAAGTATTTCCACAA 1360 |
| Qy | 1060 | GCAAGACGAGCGGGCCAAAGTCTTCAATTAAGTGAAGGACAAAGAGCTCTTCCACCATG 1119 |
| Db | 1361 | GCAAGACGAGCGGGCCAAAGTCTTCAATTAAGTGAAGGACAAAGAGCTCTTCCACCATG 1420 |
| Qy | 1120 | TGCTTCTGCTCCCTGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1179 |
| Db | 1421 | TGCTTCTGCTCCCTGT 1480 |
| Qy | 1180 | GGGGGGCTGTGTGAGACAGAGCTCCAGGACCAACCTGAGTGTATGTCTTACCTGTCTG 1239 |
| Db | 1481 | GGGGGGCTGTGTGAGACAGAGCTCCAGGACCAACCTGAGTGTATGTCTTACCTGTCTG 1540 |
| Qy | 1240 | AGTCTGATGCAACCCAAAGCGGGGCGCCGCGCTCCAGCGCCACCCACCAACAGAGAGG 1299 |
| Db | 1541 | AGTCTGATGCAACCCAAAGCGGGGCGCCGCGCTCCAGCGCCACCCACCAACAGAGAGG 1600 |
| Qy | 1300 | TTGTGCTCTTGGGGCAGATGGGCTCTGGAATCAGAAATCTTATTTGAGGACGAGAC 1359 |
| Db | 1601 | TTGTGCTCTTGGGGCAGATGGGCTCTGGAATCAGAAATCTTATTTGAGGACGAGAC 1660 |
| Qy | 1360 | CTCGGAAGCAGCGCTTAGACGGGAAAGAGCTCTCTGCTCTTCTTCAACATGAACATCTTC 1419 |
| Db | 1661 | CTCGGAAGCAGCGCTTAGACGGGAAAGAGCTCTCTGCTCTTCTTCAACATGAACATCTTC 1720 |
| Qy | 1420 | CAGAAGGACATCAACTGTGAGAGTACTACAGCTTCACTTCACTTGTAGTTTCCAGGAATTC 1479 |
| Db | 1721 | CAGAAGGACATCAACTGTGAGAGTACTACAGCTTCACTTCACTTGTAGTTTCCAGGAATTC 1780 |
| Qy | 1480 | TTTTGCAGTATGTACTATATCTTGGACGAGGGGAGGGGCGGCGGCGGCGGCGGCGGAG 1539 |
| Db | 1781 | TTTTGCAGTATGTACTATATCTTGGACGAGGGGAGGGGCGGCGGCGGCGGCGGCGGAG 1840 |
| Qy | 1540 | GTGACAGGCTGTGTGACCGGAGTACCGCTTTCTGAAAGAGGAGCTTCTTGGGACTCACGAGC 1599 |
| Db | 1841 | GTGACAGGCTGTGTGACCGGAGTACCGCTTTCTGAAAGAGGAGCTTCTTGGGACTCACGAGC 1900 |
| Qy | 1600 | CGCTTCTGTTTGGACTCTTGAACGAGGAGACCAAGGAGCCACTTGGAGAGAGTCTCTGCG 1659 |
| Db | 1901 | CGCTTCTGTTTGGACTCTTGAACGAGGAGACCAAGGAGCCACTTGGAGAGAGTCTCTGCG 1960 |
| Qy | 1660 | TGGAGGCTCTGCGGCGACATCAAGATGGACCTGTGTGAGTGGATCCAAAGCAAGAGCTCAG 1719 |
| Db | 1961 | TGGAGGCTCTGCGGCGACATCAAGATGGACCTGTGTGAGTGGATCCAAAGCAAGAGCTCAG 2020 |

QY 181 ATACCCACCTTCGGGCCAGAGGAGCGCTGGAGGTGGCTCTCAGCACCTTTGAGCGGATA 240
DB 401 ATACCCACCTTCGGGCCAGAGGAGCGCTGGAGGTGGCTCTCAGCACCTTTGAGCGGATA 460
QY 241 AACAGGAAGACCTGTGGAGAGAGGACAGAGAGGACCTGTGTAGG- 288
DB 461 AACAGGAAGACCTGTGTGGAGAGAGGACAGAGAGGACCTGTGTAGGAGATACCCACCT 520
QY 289 ----- 288
DB 521 GGTGGCCGCTCTCACTTGGGAACCAAGTCAACATGCGCTTCTGGAAGTCTCTTTGTCACT 580
QY 289 -----GATCCCGAGGAACCTACAGGAGCTATGTCCCGAGGAATTCGGGCTCATG 339
DB 581 CCAAGAAAAGATCCCGAGGAACCTACAGGAGCTATGTCCCGAGGAATTCGGGCTCATG 640
QY 340 GAACACCGCAATGCGCGCTAGGGGAATGTCTCAACCTCAGCCACCGGTACACCCGGCTC 399
DB 641 GAACACCGCAATGCGCGCTAGGGGAATGTCTCAACCTCAGCCACCGGTACACCCGGCTC 700
QY 400 CTGCTGTGTGAAGAGCACTCAAAACCCATGCAAGTCCAGCAGCAGCTTCTGGACACAGGC 459
DB 701 CTGCTGTGTGAAGAGCACTCAAAACCCATGCAAGTCCAGCAGCAGCTTCTGGACACAGGC 760
QY 460 CGGGACACCGCGAGGACCGTGGGACACAGGCTAGCCCCATCAAGATAGAGACCTCTTT 519
DB 761 CGGGACACCGCGAGGACCGTGGGACACAGGCTAGCCCCATCAAGATAGAGACCTCTTT 820
QY 520 GAGCCAGAGGAGCGCCCGAGCCACCGGACCGTGGTCAATGCAAGGCGGCGAGGG 579
DB 821 GAGCCAGAGGAGCGCCCGAGCCACCGGACCGTGGTCAATGCAAGGCGGCGAGGG 880
QY 580 ATAGGCAAGTCCATGCTGGACACAAAGGTGATGTGGACTGGGCGGACGGGAAGCTCTTC 639
DB 881 ATAGGCAAGTCCATGCTGGACACAAAGGTGATGTGGACTGGGCGGACGGGAAGCTCTTC 940
QY 640 CAAGGCAAGTTGATTTATCTCTTCAATCAATGCAAGGAGATGAACCGAGTGCCACG 699
DB 941 CAAGGCAAGTTGATTTATCTCTTCAATCAATGCAAGGAGATGAACCGAGTGCCACG 1000
QY 700 GAATGCAAGTTCGAAGACCTCATCTCAGCTGCTGGCTGAGCCGAGCGGCGCTCTCCAG 759
DB 1001 GAATGCAAGTTCGAAGACCTCATCTCAGCTGCTGGCTGAGCCGAGCGGCGCTCTCCAG 1060
QY 760 GAGCTCATCCAGTTCGAGCGGCTCTCTTTTCAATCAATGCAAGGCTTCGATGAGCTCAAG 819
DB 1061 GAGCTCATCCAGTTCGAGCGGCTCTCTTTTCAATCAATGCAAGGCTTCGATGAGCTCAAG 1120
QY 820 CTTCTTTTCCAGATCTCAGGGAACCTGTGCTCTGCTGGGAGGAGAAACGGCCCAAG 879
DB 1121 CTTCTTTTCCAGATCTCAGGGAACCTGTGCTCTGCTGGGAGGAGAAACGGCCCAAG 1180
QY 880 GAGCTGCTTTAAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATC 939
DB 1181 GAGCTGCTTTAAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATC 1240
QY 940 ACCACAGGCGCACCGCTTTGGAGAAGCTCCACCGTCTGCTGGAGCACCCAGGCATGTG 999
DB 1241 ACCACAGGCGCACCGCTTTGGAGAAGCTCCACCGTCTGCTGGAGCACCCAGGCATGTG 1300
QY 1000 GAGATCTTGGGCTTCTCTGAGGCAAGAGGAATACTTTCTAAGATATTTCCACAAT 1059
DB 1301 GAGATCTTGGGCTTCTCTGAGGCAAGAGGAAGTAATCTTCAAGATATTTCCACAAT 1360
QY 1060 GCAGAGAGGCGGCGCAAGTCTCAATTAAGTGAAGGACACAGGCTCTCTTCAACCATG 1119
DB 1361 GCAGAGAGGCGGCGCAAGTCTCAATTAAGTGAAGGACACAGGCTCTCTTCAACCATG 1420
QY 1120 TGTCTTGGTCCCTGTGGTGTGCTGAGTGTGTAACCTGCTCCAGCAGCAGCTGAGGGT 1179
DB 1421 TGTCTTGGTCCCTGTGGTGTGCTGAGTGTGTAACCTGCTCCAGCAGCAGCTGAGGGT 1480
QY 1180 GGGGGGCTGTTGAGACAGACGTCAGGACCAACCTGCAAGTGTATGCTCTACCTGCTG 1239

DB 1481 GGGGGGCTGTTGAGACAGACGTCAGGACCAACCACTGAGTGTATGCTCTACCTGCTG 1540
QY 1240 AGTCTGATGCAACCAAGCGGGGGCCCGCGCTCCAGCGCCCAACCCAGACAGAGGG 1299
DB 1541 AGTCTGATGCAACCAAGCGGGGGCCCGCGCTCCAGCGCCCAACCCAGACAGAGGG 1600
QY 1300 TTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAATCTCTATTGAGGAGCAGAC 1359
DB 1601 TTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAATCTCTATTGAGGAGCAGAC 1660
QY 1360 CTCGGAGACCGGCTTAGACGGGAAGACGTCTCTGCTCTCTCAACATGAATCTTCT 1419
DB 1661 CTCGGAGACCGGCTTAGACGGGAAGACGTCTCTGCTCTCTCAACATGAATCTTCT 1720
QY 1420 CAGAAGGACATCAACTGTGAGAGTACTACAGCTTCTCACTTGAAGTTCACAGAAATC 1479
DB 1721 CAGAAGGACATCAACTGTGAGAGTACTACAGCTTCTCACTTGAAGTTCACAGAAATC 1780
QY 1480 TTTGCACTATGTAATAATCTTGGACGAGGGGAGGGCGGGGAGGCCCAAGACAGAC 1539
DB 1781 TTTGCACTATGTAATAATCTTGGACGAGGGGAGGGCGGGGAGGCCCAAGACAGAC 1840
QY 1540 GTGACAGGCTGTGACCGAGTACCGGTTTTCTGAAAAGGAGCTTCTTGGCACTCACGAC 1599
DB 1841 GTGACAGGCTGTGACCGAGTACCGGTTTTCTGAAAAGGAGCTTCTTGGCACTCACGAC 1900
QY 1600 CGCTTCTGTTTGGACTCTTGAACGAGGAGACAGGAGCCACCTGGAGAGAGTCTCTGC 1659
DB 1901 CGCTTCTGTTTGGACTCTTGAACGAGGAGACAGGAGCCACCTGGAGAGAGTCTCTGC 1960
QY 1660 TGGAGGCTCTCGCGCACATCAAGATGGACCTGTGTGAGTGGATCCAAAGCAAGCTCAG 1719
DB 1961 TGGAGGCTCTCGCGCACATCAAGATGGACCTGTGTGAGTGGATCCAAAGCAAGCTCAG 2020
QY 1720 AGCGACGGCTCCACCTTGACGAGGCTCTTGGAGTTCCTTCACTGCTGTGTAGAGATC 1779
DB 2021 AGCGACGGCTCCACCTTGACGAGGCTCTTGGAGTTCCTTCACTGCTGTGTAGAGATC 2080
QY 1780 CAGGAGGAGGTTTTATCCAGAGGCTCTGAGCCACTTCCAGGTGATCTGAGTGTGATCAGCAAC 1839
DB 2081 CAGGAGGAGGTTTTATCCAGAGGCTCTGAGCCACTTCCAGGTGATCTGAGTGTGATCAGCAAC 2140
QY 1840 ATTGCTCTCAAGATGGAGCATGGTCTCTCTGTTCTGTCTGAAGCGCTGCAGAGCGCC 1899
DB 2141 ATTGCTCTCAAGATGGAGCATGGTCTCTCTGTTCTGTCTGAAGCGCTGCAGAGCGCC 2200
QY 1900 CAGTGTCTGACTTGTATGCGCCACCTACAGCGCGGACGGGGAAGACCCGCGCAGGTGC 1959
DB 2201 CAGTGTCTGACTTGTATGCGCCACCTACAGCGCGGACGGGGAAGACCCGCGCAGGTGC 2260
QY 1960 TCCGACGAGCGCACACGCTGTTGGTGCAGCTCAGACAGAGAGCGGTTCTGCTGCGAC 2019
DB 2261 TCCGACGAGCGCACACGCTGTTGGTGCAGCTCAGACAGAGAGCGGTTCTGCTGCGAC 2320
QY 2020 GCCTACAGTGAACATCTGGCAGCGGCTCTGCAACCAATCCAAACCTGATAGAGCTGTCT 2079
DB 2321 GCCTACAGTGAACATCTGGCAGCGGCTCTGCAACCAATCCAAACCTGATAGAGCTGTCT 2380
QY 2080 CTGTAACGAAATGCTTGGGAGCGCGGGGGTGAAGCTGTCTGTCAAGGATCAGACAC 2139
DB 2381 CTGTACCGAAATGCTTGGGAGCGCGGGGGTGAAGCTGTCTGTCAAGGATCAGACAC 2440
QY 2140 CCCAACTGCAAACTTCAAGAACCTGAGGCTGAAGAGGTGCCGCATCTCCAGCTCAGCTGC 2199
DB 2441 CCCAACTGCAAACTTCAAGAACCTGAGGCTGAAGAGGTGCCGCATCTCCAGCTCAGCTGC 2500
QY 2200 GAGGACCTCTGAGCTCTCATAGCCAAATAAGAAATTTGAAGAGATGGATCTCAGTGGC 2259
DB 2501 GAGGACCTCTGAGCTCTCATAGCCAAATAAGAAATTTGAAGAGATGGATCTCAGTGGC 2560
QY 2260 AACGGGCTTGGATTTCCAGGATGATGCTGCTTTGCGAGGCTGCGGCAATCCCCAGTGC 2319

| | | | | |
|------------|---|---|--------|--------------------------------------|
| Db | 2561 | AA CGCGTTGGATTCCAGGCATGATGCTGCTTTGGAGGGCTGCGGCATCCCAAGTGC | 2620 | number of bases overlapped is 22832. |
| Qy | 2320 | AGGCTGAGATGATGATGAGGAAGTGTACAGTCCGGGCTGTGTCAGGAGATG | 2379 | Location/Qualifiers |
| Db | 2621 | AGGCTGAGATGATGATGAGGAAGTGTACAGTCCGGGCTGTGTCAGGAGATG | 2680 | 1. 147330 |
| Qy | 2380 | GCTTCTGTGCTCGGCACCAACCATCTGTTGAGTTGACCTGACAGAAATGCATG | 2439 | /organism="Homo sapiens" |
| Db | 2681 | GCTTCTGTGCTCGGCACCAACCATCTGTTGAGTTGACCTGACAGAAATGCATG | 2740 | /mol_type="genomic DNA" |
| Qy | 2440 | GAGGATTTGGGCTGAGTTACTATGACAGGACTGAGGACCCAGTCTGACAGTACGG | 2499 | /db_xref="taxon:9606" |
| Db | 2741 | GAGGATTTGGGCTGAGTTACTATGACAGGACTGAGGACCCAGTCTGACAGTACGG | 2800 | /chromosome="19" |
| Qy | 2500 | ACTTTGTGG | 2508 | /clone="CTD-3022G6" |
| Db | 2801 | ACTTTGTGG | 2809 | |
| RESULT 12 | | | | |
| AC008753 | 147330 bp | DNA | linear | PRI 22-MAR-2003 |
| LOCUS | | | | |
| DEFINITION | Homo sapiens chromosome 19 clone CTD-3022G6, complete sequence. | | | |
| AC008753 | | | | |
| VERSION | AC008753.9 Gi:25140998 | | | |
| KEYWORDS | HTG. | | | |
| SOURCE | Homo sapiens (human) | | | |
| ORGANISM | Homo sapiens | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| TITLE | 1 (bases 1 to 147330) | | | |
| JOURNAL | Direct Submission | | | |
| AUTHORS | DOE Joint Genome Institute and Stanford Human Genome Center. | | | |
| TITLE | Unpublished | | | |
| JOURNAL | 2 (bases 1 to 147330) | | | |
| AUTHORS | DOE Joint Genome Institute. | | | |
| TITLE | Direct Submission | | | |
| JOURNAL | Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint | | | |
| AUTHORS | Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA | | | |
| TITLE | 3 (bases 1 to 147330) | | | |
| JOURNAL | DOE Joint Genome Institute and Stanford Human Genome Center. | | | |
| AUTHORS | Direct Submission | | | |
| TITLE | Submitted (29-AUG-2000) DOE Joint Genome Institute, 2800 Mitchell | | | |
| JOURNAL | Drive, Walnut Creek, CA 94598, USA | | | |
| AUTHORS | 4 (bases 1 to 147330) | | | |
| TITLE | DOE Joint Genome Institute and Stanford Human Genome Center. | | | |
| JOURNAL | Submitted (06-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell | | | |
| AUTHORS | Drive, Walnut Creek, CA 94598, USA | | | |
| TITLE | 5 (bases 1 to 147330) | | | |
| JOURNAL | DOE Joint Genome Institute and Stanford Human Genome Center. | | | |
| AUTHORS | Direct Submission | | | |
| TITLE | Submitted (21-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell | | | |
| JOURNAL | Drive, Walnut Creek, CA 94598, USA | | | |
| AUTHORS | 6 (bases 1 to 147330) | | | |
| TITLE | DOE Joint Genome Institute and Stanford Human Genome Center. | | | |
| JOURNAL | Submitted (22-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell | | | |
| AUTHORS | Drive, Walnut Creek, CA 94598, USA | | | |
| TITLE | 7 (bases 1 to 147330) | | | |
| JOURNAL | DOE Joint Genome Institute and Stanford Human Genome Center. | | | |
| AUTHORS | Direct Submission | | | |
| TITLE | Submitted (22-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell | | | |
| JOURNAL | Drive, Walnut Creek, CA 94598, USA | | | |
| AUTHORS | On Nov 21, 2002 this sequence version replaced gi:9937751. | | | |
| TITLE | Draft Sequence Produced by DOE Joint Genome Institute | | | |
| JOURNAL | www.jgi.doe.gov | | | |
| AUTHORS | Finishing Completed at Stanford Human Genome Center | | | |
| TITLE | www-shgc.stanford.edu | | | |
| JOURNAL | Quality: Phrap Quality >=40 99.6% of Sequence; | | | |
| COMMENT | Estimated Total Number of Errors is 0.5. | | | |
| | NOTE: This insert is not the entire sequence of the clone (entire | | | |
| | sequence is 160kb). It is clipped at the overlap with AC008440. The | | | |

Qy 1189 TTGAGACAGAGCTCCAGACCACTCCAGCTCCAGCCCCCACCACCAACAGAGAGGGTTGTCTCC 1248
Db 79121 TTGAGACAGAGCTCCAGACCACTCCAGCTCCAGCTGATGCTCTACCTCTGAGTCTGATG 79062
Qy 1249 CAACCCAGCGGG 1308
Db 79061 CAACCCAGCGGG 79002
Qy 1309 TTGGCGGAGAGTGGGCTCTGGAATCAGAAAAATCTTATTGAGGAGCAGAGCTCCGGAG 1368
Db 79001 TTGGCGGAGAGTGGGCTCTGGAATCAGAAAAATCTTATTGAGGAGCAGAGCTCCGGAG 78942
Qy 1369 CAGGCGCTAGACGGGGAAGAGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1428
Db 78941 CAGGCGCTAGACGGGGAAGAGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 78982
Qy 1429 ATCAACTGTGAGAGTACTACAGCTTCATCCACTTGAGTTTCCAGGANTCTTTGCACT 1488
Db 78881 ATCAACTGTGAGAGTACTACAGCTTCATCCACTTGAGTTTCCAGGANTCTTTGCACT 78822
Qy 1489 ATGACTATATCTTGGACGAGGG 1548
Db 78821 ATGACTATATCTTGGACGAGGG 78762
Qy 1549 CTGTTGACCGAGTACGCGTTTCTGAAAGAGCTTCTTGGCACTCACCAGCCGCTTCTG 1608
Db 78761 CTGTTGACCGAGTACGCGTTTCTGAAAGAGCTTCTTGGCACTCACCAGCCGCTTCTG 78702
Qy 1609 TTTGGACTCTTGAACGAGGAGACAGGAGCCACTTGGAGAGAGTCTCTGCTGAGAGTC 1668
Db 78701 TTTGGACTCTTGAACGAGGAGACAGGAGCCACTTGGAGAGAGTCTCTGCTGAGAGTC 78642
Qy 1669 TCGCGGCACATCAAGATGAGCTGTTGAGTGGATCCAAAGCAAGCTCAGAGCAGCGC 1728
Db 78641 TCGCGGCACATCAAGATGAGCTGTTGAGTGGATCCAAAGCAAGCTCAGAGCAGCGC 78582
Qy 1729 TCCACCTTGACGAGGGCTCTTGGAGTTCTTCAAGTCTTCAAGTCTGTTGAGATCAGAGAGG 1788
Db 78581 TCCACCTTGACGAGGGCTCTTGGAGTTCTTCAAGTCTTCAAGTCTGTTGAGATCAGAGAGG 78522
Qy 1789 GAGTTATCCAGAGGCGCTCAGCACTTCCAGTGTGCTGTTGAGTCTGTTGAGTCTGTTGAGTCT 1848
Db 78521 GAGTTATCCAGAGGCGCTCAGCACTTCCAGTGTGCTGTTGAGTCTGTTGAGTCTGTTGAGTCT 78462
Qy 1849 AAGATGAGACATGCTCTCTGTTCTGTTGAGTCTGTTGAGTCTGTTGAGTCTGTTGAGTCT 1908
Db 78461 AAGATGAGACATGCTCTCTGTTCTGTTGAGTCTGTTGAGTCTGTTGAGTCTGTTGAGTCT 78402
Qy 1909 CACTTGTATGCGGCACTCAGCGCGGACCGGGAAGACCGCGAGTGTCTCCGAGGA 1968
Db 78401 CACTTGTATGCGGCACTCAGCGCGGACCGGGAAGACCGCGAGTGTCTCCGAGGA 78342
Qy 1969 GCGCACAGCTGTTGGTGCAGCT 1991
Db 78341 GCGCACAGCTGTTGGTGCAGCT 78319

RESULT 13
AX575503
LOCUS AX575503 2158 bp DNA linear PAT 07-JAN-2003
DEFINITION Sequence 19 from Patent WO02072630.
ACCESSION AX575503
VERSION AX575503.1 GI:27552095
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Thornton.M., Hafalia,A.J., Lu,D.A., Arvizu,C., Swarnakar,A., Lu,Y.,
Warren,B.A., Baughn,M.R., Tang,Y.T., Lee,E.A., Yao,M.G.,
Ranikumar,J., Khan,P.A., Gandhi,A.R., Ding,L., Yue,H., Gietzen,K.J.,

Walia,N.K., Thangavelu,K., Elliot,V.S. and Marquis,J.P.
Nucleic acid-associated proteins
Patent: WO 02072630-A 19 19-SEP-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
1..2158
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 429930CB1"

Query Match 54.1%; Score 1681.8; DB 6; Length 2158;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1696; Conservative 0; Mismatches 2; Indels 3; Gaps 1;
ORIGIN
Qy 1408 ATGAACATCTTCCAGAGGACATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGT 1467
Db 1 ATGAACATCTTCCAGAGGACATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGT 60
Qy 1468 TTCCAGGAATTTTTCAGAGCTTATCTATATCTTGGACGAGGGGGGGGGGGGGGGGGGGGG 1527
Db 61 TTCCAGGAATTTTTCAGAGCTTATCTATATCTTGGACGAGGGGGGGGGGGGGGGGGGGGG 120
Qy 1528 CCAGACAGAGCTGACAGAGCTTTCAGCGAGTACCGGTTTCTGAAAGAGCTTCTG 1587
Db 121 CCAGACAGAGCTGACAGAGCTTTCAGCGAGTACCGGTTTCTGAAAGAGCTTCTG 180
Qy 1588 GCACCTACAGCGCTTCTGTTGGACTCTTGAACGAGGAGACAGGAGCAGCTGAG 1647
Db 181 GCACCTACAGCGCTTCTGTTGGACTCTTGAACGAGGAGACAGGAGCAGCTGAG 240
Qy 1648 AAGAGTCTCTGCTGGAAGTCTCGCCGCACATCAAGATGGAACCTGTTGCACTGGATCCAA 1707
Db 241 AAGAGTCTCTGCTGGAAGTCTCGCCGCACATCAAGATGGAACCTGTTGCACTGGATCCAA 300
Qy 1708 AGCAAGCTCAGAGCGAGCTCCACCTCGACAGAGGCTCTTGGAGTCTTTCAGTGC 1767
Db 301 AGCAAGCTCAGAGCGAGCTCCACCTCGACAGAGGCTCTTGGAGTCTTTCAGTGC 360
Qy 1768 TTGTACAGATCCAGAGGAGGAGTTCATCCAGCAGGCGCTTCCAGCTTCAGTGCATC 1827
Db 361 TTGTACAGATCCAGAGGAGGAGTTCATCCAGCAGGCGCTTCCAGCTTCAGTGCATC 420
Qy 1828 GTGTGTAGCAACATTTGCTTCAAGATGGAGCAGATGCTCTGTTCTGTTCTGTTCTGTTCTG 1887
Db 421 GTGTGTAGCAACATTTGCTTCAAGATGGAGCAGATGCTCTGTTCTGTTCTGTTCTGTTCTG 480
Qy 1888 TGCAAGAGCGCCAGGTGCTGCACTTGTATGGCGCCACCTACAGCGGAGCGGGAAGAC 1947
Db 481 TGCAAGAGCGCCAGGTGCTGCACTTGTATGGCGCCACCTACAGCGGAGCGGGAAGAC 540
Qy 1948 CGCGGAGGTGCTCCGAGGAGCGCACAGCTGTTGGTGGAGCTCAGACAGAGAGGAC 2007
Db 541 CGCGGAGGTGCTCCGAGGAGCGCACAGCTGTTGGTGGAGCTCAGACAGAGAGGAC 597
Qy 2008 GTTCTGTGAGCGCTTACAGTGAACATCTGCGAGCGGCGCTTGTGCACCAATCCAAACCTG 2067
Db 598 GTTCTGTGAGCGCTTACAGTGAACATCTGCGAGCGGCGCTTGTGCACCAATCCAAACCTG 657
Qy 2068 ATAGAGTGTCTTGTATCCGAAATGCGCTGGGAGCGGGGGGTGAAGTCTCTGTCAA 2127
Db 658 ATAGAGTGTCTTGTATCCGAAATGCGCTGGGAGCGGGGGGTGAAGTCTCTGTCAA 717
Qy 2128 GGACTCAGACACCCCAACTTCAGAACTTCAGAACTTCAGAACTTCAGAACTTCAGAACTTC 2187
Db 718 GGACTCAGACACCCCAACTTCAGAACTTCAGAACTTCAGAACTTCAGAACTTCAGAACTTC 777
Qy 2188 AGCTCAGCTGCGAGGAGCTCTCTGAGCTCTCTAGCCAAATAAGAAATTTGCAAGGATG 2247
Db 778 AGCTCAGCTGCGAGGAGCTCTCTGAGCTCTCTAGCCAAATAAGAAATTTGCAAGGATG 837
Qy 2248 GATCTCAGTGGCAACCGCGGTTGGATTCCAGGCAATGATGCTGCTTTGCGAGGGCTCGCG 2307

| | | | |
|----|------|---|------|
| Db | 838 | GATCTCAGTGGCAACGGCGTGGATCCAGGCGATGATCTGCTTTGCGAGGCGCTGCGG | 897 |
| Qy | 2308 | CATCCCAAGTGCAGGCTGCAGATGATTCAGTTGAGGAAGTGCAGTGGAGTCCGGGCT | 2367 |
| Db | 898 | CATCCCAATGCAAGGCTGCAGATGATTCAGTTGAGGAAGTGCAGTGGAGTCCGGGCT | 957 |
| Qy | 2368 | TGTCAGGAGATGGCTTCTGCTGCTGCGGACCAACCCACATCTGGTTGAGTTGACCTTGACA | 2427 |
| Db | 958 | TGTCAGGAGATGGCTTCTGCTGCTGCGGACCAACCCACATCTGGTTGAGTTGACCTTGACA | 1017 |
| Qy | 2428 | GGAAATGCACTGAGGAGATTTGGCGCTGAGTTACTATGCGAGGAGTGGAGGACCCAGTC | 2487 |
| Db | 1018 | GGAAATGCACTGAGGAGATTTGGCGCTGAGTTACTATGCGAGGAGTGGAGGACCCAGTC | 1077 |
| Qy | 2488 | TGCAGACTGAGGACTTTGGCTGAGATCTGCGGCTCACTGCTGCTGCTGTCGAG | 2547 |
| Db | 1078 | TGCAGACTGAGGACTTTGGCTGAGATCTGCGGCTCACTGCTGCTGCTGTCGAG | 1137 |
| Qy | 2548 | CTGGCCTCAACTCTCAGTGTGAACAGAGCCTGAGAGAGTGGACCTGAGCCTGAATGAG | 2607 |
| Db | 1138 | CTGGCCTCAACTCTCAGTGTGAACAGAGCCTGAGAGAGTGGACCTGAATGAG | 1197 |
| Qy | 2608 | CTGGGAGACTCGGGGTGCTGCTGCTGTGAGGGCTCAGGCATCCAGTGCAGAGCTC | 2667 |
| Db | 1198 | CTGGGAGACTCGGGGTGCTGCTGCTGTGAGGGCTCAGGCATCCAGTGCAGAGCTC | 1257 |
| Qy | 2668 | CAGACCTCGGGTGGGCTGCTGCGGCTGAGGCTCTGCGCTGTGAGGGTCTTTCTGTG | 2727 |
| Db | 1258 | CAGACCTCGGGTGGGCTGCTGCGGCTGAGGCTCTGCGCTGTGAGGGTCTTTCTGTG | 1317 |
| Qy | 2728 | GTGCTCCAGGCAACCAACTCGGGAGCTGAGTGTGAGTTTCAAGCACTGGGAGAC | 2787 |
| Db | 1318 | GTGCTCCAGGCAACCAACTCGGGAGCTGAGTGTGAGTTTCAAGCACTGGGAGAC | 1377 |
| Qy | 2788 | TGGGCGCTGTGCTGCTGAGGGCTGCAACATCCCGCTGCAGATCCAGAACTG | 2847 |
| Db | 1378 | TGGGCGCTGTGCTGCTGAGGGCTGCAACATCCCGCTGCAGATCCAGAACTG | 1437 |
| Qy | 2848 | TGGCTGGATAGTGTGGCTCAGCAGCAAGGCTGTGAGATCTTTACTTCACTCCCTGGG | 2907 |
| Db | 1438 | TGGCTGGATAGTGTGGCTCAGCAGCAAGGCTGTGAGATCTTTACTTCACTCCCTGGG | 1497 |
| Qy | 2908 | ATCAACCAAGACTTCAACGACCTTTACTGACCAACGCTTGGAGGACACAGGTGTC | 2967 |
| Db | 1498 | ATCAACCAAGACTTCAACGACCTTTACTGACCAACGCTTGGAGGACACAGGTGTC | 1557 |
| Qy | 2968 | CGACTGCTTTGCAAGCGCTGAGCCATCTGGCTGCAAACTCCGAGTCTCTGTTATT | 3027 |
| Db | 1558 | CGACTGCTTTGCAAGCGCTGAGCCATCTGGCTGCAAACTCCGAGTCTCTGTTATT | 1617 |
| Qy | 3028 | GGATGGACCTGAATTAATGACCAAGTGTGGAGGCTTGGAGTAAACAACT | 3087 |
| Db | 1618 | GGATGGACCTGAATTAATGACCAAGTGTGGAGGCTTGGAGTAAACAACT | 1677 |
| Qy | 3088 | TATTTGACATTTGGCTGCTGA | 3108 |
| Db | 1678 | TATTTGACATTTGGCTGCTGA | 1698 |

RESULT 14
AX684291
LOCUS
DEFINITION
Sequence 1 from Patent WO02052011.
AX684291
ACCESSION
AX684291.1
VERSION
GI:29371161
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Feder, J., Ramanathan, C. and Mintier, G.

| | |
|----------|---|
| TITLE | Human leucine-rich repeat containing protein, hlrblm1, expressed predominantly in bone marrow |
| JOURNAL | Patent: WO 02052011-A 1 04-JUL-2002; |
| FEATURES | Bristol-Myers Squibb Company (US) |
| source | Location/Qualifiers |
| | 1..4931 |
| | /organism="Homo sapiens" |
| | /mol_type="unassigned DNA" |
| | /db_xref="taxon:9606" |
| CDS | 1320..2669 |
| | /note="unnamed protein product" |
| | /codon_start=1 |
| | /protein_id="CAD86517.1" |
| | /db_xref="GI:29371162" |
| | /translation="MNQATECSMDLIFSCWPEPSAPLOELIRVPERLLFIIDPDE LKPSFHPDQGPWCLCWEKPEPTLLNSLLKLLPELSLITTRPTALEKHLLEH PRHVEILGFSEAEKKEYFYKYNHAEQAGOVFNVRDNEPLFTWCFVPLVWVVTCL QOQLEGGLLKQTSRTTAYMLYLLSLMQPKAPLQPPNORGLCSLAADGLMNO KILFEQDLRKHLGDSDVSAFLNNMTFQKIDINCERSFLATSRFLGLNEETRSH EKSLCWKVSPIHKMDLQWIOSKAQSDGSTLQGSLEFFSCFLYIQRBEFTIQALSHF QVIVVSNIAKMEHWSFCLKRCRQAQVHLGYATYSDGDRARCSAGHTLLVLQ PERTVLLDAYSEHLAAALCTNPNLIELSLYRNALSGRGVKKLQGLRHPNCKLQNLK |
| ORIGIN | |
| | Query Match 52.0%; Score 1616; DB 6; Length 4931; |
| | Best Local Similarity 81.0%; Pred. No. 0; |
| | Matches 2160; Conservative 0; Mismatches 5; Indels 501; Gaps 4; |
| Qy | 289 GATCCCCAGGAACTCAGGAGCTATGTCGCGAGGAAATTCGGCTCATCGAAGACGC 348 |
| Db | 1021 GATCCCCAGGAACTCAGGAGCTATGTCGCGAGGAAATTCGGCTCATCGAAGACGC 1080 |
| Qy | 349 AATGGCGCTAGGGGAATGTCAACCTCAGGCAACCGGTACACCGGCTCTCTGCTGTG 408 |
| Db | 1081 AATGGCGCTAGGGGAATGTGTCAACCTCAGGCAACCGGTACACCGGCTCTCTGCTGTG 1140 |
| Qy | 409 AAGGAGCACTCAACCCATGCAGCTCAGCAGAGCTTCTTGACACAGCGCGGAGCAC 468 |
| Db | 1141 AAGGAGCACTCAACCCATGCAGCTCAGCAGAGCTTCTTGACACAGCGCGGAGCAC 1200 |
| Qy | 469 GCGAGACCGTGGGACACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCAGAC 528 |
| Db | 1201 GCGAGACCGTGGGACACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCAGAC 1260 |
| Qy | 529 GAGGAGCGCGGAGCACCGGACCGGTGCTCATGCAAGCGCGCGAGGATAGCAAG 588 |
| Db | 1261 GAGGAGCGCGGAGCACCGGACCGGTGCTCATGCAAGCGCGCGAGGAT----- 1313 |
| Qy | 589 TCCATGCTGSCACACAAGGATGATGCTGAGCTGGCGGACGGGAAAGCTTTTCCAGGACGA 648 |
| Db | 1314 ----- 1313 |
| Qy | 649 TTTGATTATCTTTCTTACATCAACTCAGGAGATGAACAGAGTGCCACGGAATGCAGC 708 |
| Db | 1314 -----AGGAGATGAACAGAGTGCCACGGAATGCAGC 1346 |
| Qy | 709 ATGCAAGACCTCATCTTCTGCTGCTGAGCCCGAGCGCGCTCTCCAGGAGCTCATC 768 |
| Db | 1347 ATGCAAGACCTCATCTTCTGCTGCTGAGCCCGAGCGCGCTCTCCAGGAGCTCATC 1406 |
| Qy | 769 CGAGTTCCCGAGCGCTCTTTTTCATCATCGACGGCTTCGATGAGCTCAAGCCTTCTTTC 828 |
| Db | 1407 CGAGTTCCCGAGCGCTCTTTTTCATCATCGACGGCTTCGATGAGCTCAAGCCTTCTTTC 1466 |
| Qy | 829 CACATCTCAGGAGACCTGCTGCTGCTGGAGGAGAAACCGGCCACCGAGCTGCTT 888 |
| Db | 1467 CACATCTCAGGAGACCTGCTGCTGCTGGAGGAGAAACCGGCCACCGAGCTGCTT 1526 |
| Qy | 889 CTTAACAGCTTAATTCGAAAGAGCTGCTCTCTGAGCTATCTTTGCTCATCAACACGG 948 |
| Db | 1527 CTTAACAGCTTAATTCGAAAGAGCTGCTCTCTGAGCTATCTTTGCTCATCAACACAG 1586 |
| Qy | 949 CCCACGGCTTTGGAGAGCTCCACCGCTCTGCTGAGGACCCCGAGCATGTGGAGATCTG 1008 |

Db 1587 |||||CCCCCGGCTTTGGAGAGCTCCACCGTCTGCTGGAGCACCACCGCATGTGGAGATCGT 1646
Qy GCTTCTCTGAGGAGAGAGGAATACCTTCTACAGATATTTCACAAATGACAGAGCAG 1068
Db 1647 GCTTCTCTGAGGAGAGAGAGGAATACCTTCTACAGATATTTCACAAATGACAGAGCAG 1706
Qy 1069 GCGGGCCAAAGCTTTCAATTAAGTGGAGGACCAACGAGCCCTCTCTTCCACCATGTGCTTCGTC 1128
Db 1707 GCGGGCCAAAGCTTTCAATTAAGTGGAGGACCAACGAGCCCTCTCTTCCACCATGTGCTTCGTC 1766
Qy 1129 CCCCTGGTGTGCTGGTGGTGTGTACCTGCTCCAGCAGCAGCTGGAGGGTGGGGGGCTG 1188
Db 1767 CCCCTGGTGTGCTGGTGGTGTGTACCTGCTCCAGCAGCAGCTGGAGGGTGGGGGGCTG 1826
Qy 1189 TTGAGACAGAGCTCCAGAGCACCACCTGACGTGTACATGCTCTACCTGCTGAGTCTGATG 1248
Db 1827 TTGAGACAGAGCTCCAGAGCACCACCTGACGTGTACATGCTCTACCTGCTGAGTCTGATG 1886
Qy 1249 CAACCCAGCGGGGGCCCCCGGCTCCAGCCCCCACCACCAACAGAGAGGGTTGTGCTCC 1308
Db 1887 CAACCCAGCGGGGGCCCCCGGCTCCAGCCCCCACCACCAACAGAGAGGGTTGTGCTCC 1946
Qy 1309 TTGGGGCAGATGGGCTCTGAAATCAGAAATCCTATTGAGGAGCAGGACCTCCGGAAG 1368
Db 1947 TTGGGGCAGATGGGCTCTGAAATCAGAAATCCTATTGAGGAGCAGGACCTCCGGAAG 2006
Qy 1369 CACGGCTTAGACGGGAGAGCTCTGCTCTCTCAACATGAACATCTTCCAGAGAGGAC 1428
Db 2007 CACGGCTTAGACGGGAGAGAGCTCTGCTCTCTCAACATGAACATCTTCCAGAGAGGAC 2066
Qy 1429 ATCAACTGTGAGAGGACTACAGCTTCACTCACTTTGAGTTTCCAGGAATTTCTTGGCAGCT 1488
Db 2067 ATCAACTGTGAG----- 2078
Qy 1489 ATGTAATATATCTTGACAGAGGGGAGGGCGGGGAGGCCACAGACCAGGAGCTGACCAGG 1548
Db 2079 ----- 2078
Qy 1549 CTGTTGACCGAGTAGCGGTTTCTGAAAGAGCTTCTTGACACTCACCGACCGCTTCTG 1608
Db 2079 -----AGGAGCTTCTTGGCAGCTCACCGCCGCTTCTG 2111
Qy 1609 TTTGGACTCTGAACGAGGAGACAGGAGCACCCTGGAGAGAGTCTCTGCTGGAAGTC 1668
Db 2112 TTTGGACTCTGAACGAGGAGACAGGAGCACCCTGGAGAGAGTCTCTGCTGGAAGTC 2171
Qy 1669 TCGCCGCACATCAAGATGGACCTGTTGCAAGTGATCCAAAGCAAAAGCTCAGAGCGACGGC 1728
Db 2172 TCGCCGCACATCAAGATGGACCTGTTGCAAGTGATCCAAAGCAAAAGCTCAGAGCGACGGC 2231
Qy 1729 TCCACCTGACAGCGGCTCTTGAGTTCTTGAGTCTTACGCTGTTGACGAGATCCAGAGGAG 1788
Db 2232 TCCACCTGACAGCGGCTCTTGAGTTCTTGAGTCTTACGCTGTTGACGAGATCCAGAGGAG 2291
Qy 1789 GAGTTTATCCAGCAGGCTTCCAGCACAATTCCAGGTGATCGTGGTCAAGCAATTCGCTCC 1848
Db 2292 GAGTTTATCCAGCAGGCTTCCAGCACAATTCCAGGTGATCGTGGTCAAGCAATTCGCTCC 2351
Qy 1849 AAGATGAGACATGCTCTCTGTTCTGTTCTGTAAGCGCTCAGAGAGGCCACAGGTGCTG 1908
Db 2352 AAGATGAGACATGCTCTCTGTTCTGTTCTGTAAGCGCTCAGAGAGGCCACAGGTGCTG 2411
Qy 1909 CACTTGTATGGCCACCTACAGCGGAGCGGGAGACCGCGGAGGTGCTCCGACGGA 1968
Db 2412 CACTTGTATGGCCACCTACAGCGGAGCGGGAGACCGCGGAGGTGCTCCGACGGA 2471
Qy 1969 GCGCACAGCTGTTGGTGCAGCTCAGACAGAGAGGACCGTTCCTGCTGAGCGCTTACAGT 2028
Db 2472 GCGCACAGCTGTTGGTGCAGCT---ACAGAGAGGACCGTTCCTGCTGAGCGCTTACAGT 2528
Qy 2029 GAACATCTGGAGCGGCCCTGTGCAACCAATCCAACTGTGATGAGCTGCTCTGTGACCGA 2088
|||||

Db 2529 GAACATCTGGACGCGGCCCTGTGCAACCAATCCAAACCTGATAGAGCTGTCTCTGTACCGA 2588
Qy 2089 AATGCCCTCGGACGCGGGGGTGAAGCTGTCTGTCAAGAGCTCAGACACCCCACTGC 2148
Db 2589 AATGCCCTCGGACGCGGGGGTGAAGCTGTCTGTCAAGAGCTCAGACACCCCACTGC 2648
Qy 2149 AAATCTCAGAACCTGA----- 2164
Db 2649 AAATCTCAGAACCTGAGGTAAATTTATCATATATAACATGATATTTTGAATAAATATA 2708
Qy 2165 ----- 2164
Db 2709 TTGGCCAGGTATGATGGCTCAGCCTGTAATTCAGCACCTTTGGGAGGCCCGCAGATGGGA 2768
Qy 2165 ----- 2164
Db 2769 GGATCACTTGACCCAGGAGTTCAAGACCGCTGGCCCAACATGGTGAAACCCCATCTCTA 2828
Qy 2165 ----- 2164
Db 2829 CTAATAATACCAAAATGAGCCAGGCATGGTGCCACAGCTCTGTAAAGCCAGCTACTCAGG 2888
Qy 2165 -----GGCTGAAGAGGTGCC 2179
Db 2889 AGGCCAAGGAGGAGGATGCTTCAACCCAGGAGGAGAGGTTGTGGCTGAAGAGGTGCC 2948
Qy 2180 GCATCTCAGCTCAGCCCTGCGAGGACCTCTCTGAGCTCTCATAGCCAAATAAGAAATTGA 2239
Db 2949 GCATCTCAGCTCAGCCCTGCGAGGACCTCTCTGAGCTCTCATAGCCAAATAAGAAATTGA 3008
Qy 2240 CAAGGATGGAATCTCAGTGGCAACGCGCTTGGAATTCAGGAGCATGCTGCTTTGCGAGG 2299
Db 3009 CAAGGATGGAATCTCAGTGGCAACGCGCTTGGAATTCAGGAGCATGCTGCTTTGCGAGG 3068
Qy 2300 GCCTGGCGCATCCCCAGCTGCGAGCTGCAGATGATTCAAGTTGAGGAAGTGTGAGCTGGAGT 2359
Db 3069 GCCTGGCGCATCCCCAATGCGAGCTGCAGATGATTCAAGTTGAGGAAGTGTGAGCTGGAGT 3128
Qy 2360 CCGGGGCTTGTGAGGAGATGGCTTCTGTGCTCGGCACCAACCCACATCTGGTTGAGTTGG 2419
Db 3129 CCGGGGCTTGTGAGGAGATGGCTTCTGTGCTTGGCACCACCCACATCTGGTTGAGTTGG 3188
Qy 2420 ACTGACAGGAAATGCACTGAGGATTTGGGCTTGAGGTTACTATGCGAGGACTGAGGC 2479
Db 3189 ACTGACAGGAAATGCACTGAGGATTTGGGCTTGAGGTTACTATGCGAGGACTGAGGC 3248
Qy 2480 ACCAGCTGCGAGACTACGAGCTTTGGCTGAGAGATCTCGGCTCACTGCTGCTGCT 2539
Db 3249 ACCAGCTGCGAGACTACGAGCTTTGGCTGAGAGATCTCGGCTCACTGCTGCTGCT 3308
Qy 2540 GTGACGAGCTGGCCCTCAACTCTCAGTGTGAACACAGAGCCTGAGAGAGCTGGACCTGAGCC 2599
Db 3309 GTGACGAGCTGGCCCTCAACTCTCAGTGTGAACACAGAGCCTGAGAGAGCTGGACCTGAGCC 3368
Qy 2600 TGAATGAGCTGGGGACCTCGGGGTGCTGCTGTGTGTGAGGGCTCAGGCAATCCACGT 2659
Db 3369 TGAATGAGCTGGGGACCTCGGGGTGCTGCTGTGTGTGAGGGCTCAGGCAATCCACGT 3428
Qy 2660 GCAGCTCCAGACCCCTGGGTTGGC 2685
Db 3429 GCAAGCTCCAGACCCCTGGGTTGATC 3454

RESULT 15

AC150815

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AC150815 202418 bp DNA linear HTG 13-OCT-2004
Callithrix jacchus clone CH259-282L8, WORKING DRAFT SEQUENCE, 19
unordered pieces.
AC150815
AC150815.2 GI:54111275
HTG; HTGS PHASE1; HTGS_DRAFT.
Callithrix jacchus (white-tufted-ear marmoset)

[illegible]

| | | | | | | | | | |
|---|--------|--|--------|--|--|--|--|--|--|
| Best Local Similarity 93.0%; Pred. No. 0; Matches 1583; Conservative 0; Mismatches 120; Indels 0; Gaps 0; | | | | | | | | | |
| QY | 289 | GATCCCCAGGAACCTACAGGACTATGTCCGACAGAAATTCGGCTCATGGAAGACCGC | 348 | | | | | | |
| Db | 135499 | GATCCCCAGGAAGCTACAGAGATTAGTAGCAGGAATTCGGCTCATGGAAGACCGC | 135558 | | | | | | |
| QY | 349 | AATGCGCCCTAGGGGAATGTCAACTCAGCCACCGGTACACCCGGCTCCTCTGGTG | 408 | | | | | | |
| Db | 135559 | AATGCGCCCTAGGGGAATGTCAACCTCAGCCACCGGTACACCCGGCTCCTCTGGTG | 135618 | | | | | | |
| QY | 409 | AAGAGCACTCAACCCCATCGAGTCCAGCAGCAGCTTCTGGACACAGCCGGGGACAC | 468 | | | | | | |
| Db | 135619 | AAGAGCACTCAACCCCATCGGGCCAGCAGCAGCTTCTGGACACAGCCGGGGACAC | 135678 | | | | | | |
| QY | 469 | GCGAGGACCGTGGGACACAGGCTAGCGCCATCAAGATAGAGACCTCTTTGAGCCAGAC | 528 | | | | | | |
| Db | 135679 | GCGAGGACCGTGGGATACCAACACAGCCCATCAAGATAGAGACCTCTTTGAGCCAGAT | 135738 | | | | | | |
| QY | 529 | GAGAGGCCCCGAGCCACCGCAACCGTGTCTATGCAAGCGCGGACGGGATAGGCAAG | 588 | | | | | | |
| Db | 135739 | GAGAGGCCCCGAGCCACCGCAACCGTGTCTATGCAAGCGCGGAGGATAGGCAAG | 135798 | | | | | | |
| QY | 589 | TCCATGCTGGCACACAAGTGATGCTGGAATGCTGGCGGACGGGAAGCTCTTCAAGGCAGG | 648 | | | | | | |
| Db | 135799 | TCCATGCTGGCTCAAGGTGATGCTGGAATGCTGGCGGACGGGAAGCTCTTCAAGGCAGG | 135858 | | | | | | |
| QY | 649 | TTTGATTATCTCTTACATCACTGACGGGAGATGAACACAGAGTGCCACGGAAATGCAGC | 708 | | | | | | |
| Db | 135859 | TTTGATTATCTCTTACATCACTGACGGGAGTGAAACACAGAGTGCCCGGAAACGAGC | 135918 | | | | | | |
| QY | 709 | ATGCAAGACCTCATCTTTCAGCTGTGGCTGAGCCCAAGCGCGCTCTCCAGGAGCTCATC | 768 | | | | | | |
| Db | 135919 | ATGCAAGACCTCATCTTCAGCTGTGGCTGAGCCCAAGCGCGCTCTCCAGGAGCTAGTG | 135978 | | | | | | |
| QY | 769 | CGAGTTCGAGCGGCTCTTTTCATCATCAAGCGCTTCGATGAGCTCAAGCCTCTTTC | 828 | | | | | | |
| Db | 135979 | CAGCTTCGAGCGCTCTCTTTTCATCATCAAGCGCTTCGATGAGCTCAAGCCTCTTTC | 136038 | | | | | | |
| QY | 829 | CACGATCTCAGGAGCCCTGTGTCTCTGCTGGAGGAGAAACCGCCACGAGCTGTT | 888 | | | | | | |
| Db | 136039 | CACGATCTCAGGAGCCCTGTGTCTCTGCTGGAGGAGAAACCGCCACGAGCTGTC | 136098 | | | | | | |
| QY | 889 | CTTAACAGCTTAATTCGGAAGAAGTGCTCCCTGAGCTATCTTTGCTCATCACACACGG | 948 | | | | | | |
| Db | 136099 | CTTAGCAGCTTAATTCGGAAGAAGTGCTCCCTGAGCTATCTTTGCTCATCACACACGG | 136158 | | | | | | |
| QY | 949 | CCCAGGCTTTGGAGAGCTCCACGCTCTGTGTGAGCACCCAGGCAATGAGATCCTG | 1008 | | | | | | |
| Db | 136159 | CCCAGGCTTTGGAGAGCTCCACGCTCTGTGTGAGCACCCAGGCAATGAGATCCTG | 136218 | | | | | | |
| QY | 1009 | GGCTTCTGAGGAGAGGAAGGAATCTTCTACAGTATTTCCAAATGCAGAGCAG | 1068 | | | | | | |
| Db | 136219 | GGCTTCTGAGGAGAGAGGAGGAGTACTTCTACAGTATTTCCAAATGCAGAGCAG | 136278 | | | | | | |
| QY | 1069 | GCGGCCCAAGCTTTCAATTACGTGAGGGAACAAGCGCTCTTTCAACCATGTGCTTCGTC | 1128 | | | | | | |
| Db | 136279 | GCGGCCCAAGCTTTCAATTACGTGAGGGAACAAGCGCTCTTTCAACCATGTGCTTCGTC | 136338 | | | | | | |
| QY | 1129 | CCCTGTGTGTGGTGGTGTGTAACCTGCTCAGCAGCAGCTGAGGGTGGGGGGCTG | 1188 | | | | | | |
| Db | 136339 | CCCTGTGTGTGGTGGTGTGTAACCTGCTCAGCAGCAGCTGAGGGTGGGGGGCTG | 136398 | | | | | | |
| QY | 1189 | TTGAGACAGAGCTCCAGGACCAACCTGACGTGTACATGCTCTACCTGCTGAGTCTGATG | 1248 | | | | | | |
| Db | 136399 | CTGAGACAGACATCCAGGACCAACCTGACGTGTACATGCTCTATCTGCTGAGTCTGATG | 136458 | | | | | | |
| QY | 1249 | CAACCCAGCGGGGGCCCCGGCTCCAGCGCCCCACCAACCAAGAGAGGGTTGTGTCC | 1308 | | | | | | |
| Db | 136459 | CAACCCAGCGGGGCAACCCCGGCTGAGCGCCCCACCAACCAAGAGAGGGTTGTGTCC | 136518 | | | | | | |
| QY | 1309 | TTGGCGGAGATGGGCTCTGGAATCAGAAATTCCTATTGAGGAGCAGGACCTCCGGAAG | 1368 | | | | | | |

Search completed: February 8, 2005, 20:14:15
Job time : 12969 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2005, 14:51:26 ; Search time 1504 Seconds
(without alignments)
12233.082 Million cell updates/sec

Title: US-10-781-294-23

Perfect score: 3108

Sequence: 1 atgtctacgaaccgaggag.....attggacattggtctgtga 3108

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|-------------|---------------------|
| 1 | 3108 | 100.0 | 3108 | 6 ADE36416 | Ade36416 Human PAA |
| 2 | 3108 | 100.0 | 3108 | 12 ADJ19338 | Adj19338 Human PAN |
| 3 | 3057 | 98.4 | 3300 | 6 AAL47129 | Aal47129 Pyrin dom |
| 4 | 3017 | 97.1 | 3731 | 12 ADP47723 | Adp47723 Human Mon |
| 5 | 2997.8 | 96.5 | 3186 | 6 AAL44363 | Aal44363 Human PYR |
| 6 | 2997.8 | 96.5 | 3186 | 10 ADF94754 | Adf94754 Human PYR |
| 7 | 2929 | 94.2 | 3218 | 12 ADJ19382 | Adj19382 Human PAN |
| 8 | 2759 | 88.8 | 3563 | 12 ADP47725 | Adp47725 Human Mon |
| 9 | 2644.2 | 85.1 | 3466 | 11 ADM02033 | Adm02033 Human cDN |
| 10 | 2470 | 79.5 | 3395 | 12 ADP47727 | Adp47727 Human Mon |
| 11 | 2449.4 | 78.8 | 3306 | 10 ADC30316 | Adc30316 Human nov |
| 12 | 2417 | 77.0 | 3221 | 12 ADP47729 | Adp47729 Human Mon |
| 13 | 1864.8 | 60.0 | 3102 | 12 ADP47731 | Adp47731 Murine Mo |
| 14 | 1705.4 | 54.9 | 1800 | 8 ACD03623 | Acd03623 Novel hum |
| 15 | 1681.8 | 54.1 | 2158 | 6 ABS78719 | Abse78719 Human cDN |
| 16 | 1680 | 54.1 | 1704 | 12 ACH91465 | Ach91465 Human gen |
| 17 | 1616 | 52.0 | 4931 | 6 ABL59333 | Ab159333 Nucleotid |
| 18 | 1491.4 | 48.0 | 1684 | 12 ADM93802 | Adm93802 DNA encod |
| 19 | 1394.8 | 44.9 | 1682 | 12 ADM93804 | Adm93804 DNA encod |
| 20 | 1394.8 | 44.9 | 1683 | 8 ACD03624 | Acd03624 Novel hum |

| | | | | | | |
|----|-------|------|------|----|----------|--------------------|
| 21 | 803.6 | 25.9 | 2835 | 12 | ADP47755 | Adp47755 Human CAT |
| 22 | 803.6 | 25.9 | 4170 | 13 | ADP24013 | Adp24013 PRO polyp |
| 23 | 803.6 | 25.9 | 4344 | 11 | ADM34349 | Adm34349 Human cry |
| 24 | 800.4 | 25.8 | 3857 | 4 | AAD14323 | Aad14323 Human PYR |
| 25 | 800.4 | 25.8 | 3857 | 9 | ABX93556 | Abx93556 Huma cDNA |
| 26 | 800.4 | 25.8 | 3857 | 9 | ACD27909 | Acd27909 Human PYR |
| 27 | 783.2 | 25.2 | 2934 | 12 | ADP47869 | Adp47869 Human CAT |
| 28 | 640.6 | 20.6 | 2154 | 12 | ADP47753 | Adp47753 Human CAT |
| 29 | 639 | 20.6 | 2524 | 6 | ADE36451 | Ade36451 Human PAA |
| 30 | 639 | 20.6 | 2524 | 12 | ADJ19373 | Adj19373 Human PAN |
| 31 | 626.6 | 20.2 | 2546 | 10 | ADF81910 | Adf81910 Leukaemia |
| 32 | 594.6 | 19.1 | 2847 | 4 | AAS01487 | Aas01487 Human sec |
| 33 | 594.6 | 19.1 | 2847 | 8 | ABZ73494 | Abz73494 Secreted |
| 34 | 594.6 | 19.1 | 2847 | 8 | ADA98038 | Ada98038 Human sec |
| 35 | 594.6 | 19.1 | 2847 | 10 | ADC20194 | Adc20194 Human sec |
| 36 | 594.6 | 19.1 | 2847 | 10 | ADF10619 | Adf10619 Human sec |
| 37 | 588.4 | 18.9 | 591 | 10 | ADC32201 | Adc32201 Human nov |
| 38 | 509 | 16.4 | 509 | 12 | ACH77762 | Ach77762 Human gen |
| 39 | 478.6 | 15.4 | 3039 | 12 | ADP47733 | Adp47733 Human inf |
| 40 | 477.8 | 15.4 | 506 | 5 | AAS68757 | Aas68757 DNA encod |
| 41 | 470.4 | 15.1 | 1557 | 6 | AAL47143 | Aal47143 Pyrin dom |
| 42 | 456 | 14.7 | 2808 | 12 | ADP47735 | Adp47735 Human inf |
| 43 | 453.4 | 14.6 | 487 | 9 | ACH36225 | Ach36225 Human end |
| 44 | 450.8 | 14.5 | 479 | 3 | AAC76566 | Aac76566 Human ORF |
| 45 | 438.4 | 14.1 | 2775 | 12 | ADO50889 | Ado50889 Human NOV |

ALIGNMENTS

RESULT 1

ADE36416

ID ADE36416 standard; cDNA; 3108 BP.

XX ADE36416;

XX 29-JAN-2004 (first entry)

XX Human PAA and nucleotide binding protein PAN6 cDNA.

XX cytosolic; immunosuppressive; vulnary; antiinflammatory; vasotropic;
XX antiallergic; antiulcer; dermatological; cerebroprotective; cardiac;
XX antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
XX NFkappaB activation inhibitor; PAA domain containing polypeptide;
XX PAA and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
XX apoptosis-associated speck-like protein; caspase recruitment domain 2;
XX ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
XX apoptosis; NFkappaB induction; cytokine processing;
XX cytokine receptor signaling caspase-mediated proteolysis;
XX c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
XX inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
XX neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
XX fibrosis; smooth muscle cell proliferation; balloon angioplasty;
XX restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
XX arthritis; lupus; schrojen's syndrome; Crohn's disease;
XX ulcerative colitis; graft versus host disease; stroke; heart failure;
XX neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
XX cancer therapy; PAA domain family; human; PAN6; gene; ss.

OS Homo sapiens.

XX US2003077699-A1.

XX 24-APR-2003.

XX 25-SEP-2001; 2001US-00965621.

XX 26-SEP-2000; 2000US-00671760.

XX 26-SEP-2000; 2000US-0367367P.

XX (REED/) REED J C.

XX (GODZ/) GODZIK A.

XX (CHUZ/) CHU Z.

| | | | |
|------|----|--|------|
| 361 | Qy | GGGGAATGTGTCAACTTACGCCACCGGTACACCGGCTCTCTGCTGGTGAAGGACACTCA | 420 |
| 361 | Db | GGGGAATGTGTCAAACTTCAGCCACCGGTACACCGGCTCTCTGCTGGTGAAGGACACTCA | 420 |
| 421 | Qy | AACCCCATGCAAGTCCAGCAGCAGCTTCTTGACACAGGCGGGGACACGCGAGACCGTG | 480 |
| 421 | Db | AACCCCATGCAAGTCCAGCAGCAGCTTCTTGACACAGCGCGGGACACGCGAGACCGTG | 480 |
| 481 | Qy | GGACACCAAGGTAGCCCCCATCAAGATAGAGACCTCTTTGAGCCAGAGGAGCGGCC | 540 |
| 481 | Db | GGACACCAAGGTAGCCCCCATCAAGATAGAGACCTCTTTGAGCCAGAGGAGCGGCC | 540 |
| 541 | Qy | GAGCCACCGCACCGTGTTCATCAAGGCGCGCAGGGATAGGCAAGTCCATGCTGGCA | 600 |
| 541 | Db | GAGCCACCGCGCACCGTGGTTCATGCAAGCGCGCGAGGATAGGCAAGTCCATGCTGGCA | 600 |
| 601 | Qy | CACAAGGTGATGCTGGACTGGGCGGACGGGAAGCTCTTCCAAGGCGAGATTGATTCTC | 660 |
| 601 | Db | CACAAGGTGATGCTGGACTGGGCGGACGGGAAGCTCTTCCAAGGCGAGATTGATTCTC | 660 |
| 661 | Qy | TTCTACATCAACTGCAGGGAGTGAACACAGAGTGCCACGGAAATGCAGCATGCAAGACCTC | 720 |
| 661 | Db | TTCTACATCAACTGCAGGGAGTGAACACAGAGTGCCACGGAAATGCAGCATGCAAGACCTC | 720 |
| 721 | Qy | ATCTTTCAAGTGTGGCTTGAGCCACAGCGCGCTCTTCCAGGAGCTCATCCGAGTTCCCGAG | 780 |
| 721 | Db | ATCTTTCAAGTGTGGCTTGAGCCACAGCGCGCTCTTCCAGGAGCTCATCCGAGTTCCCGAG | 780 |
| 781 | Qy | CGCCTCTTTTCATCATCGA CGGCTTCATGAGCTCAAGCTTCTTTCCACGATCCTCTCAG | 840 |
| 781 | Db | CGCCTCTCTTTTCATCATCGA CGGCTTCATGAGCTCAAGCTTCTTTCCACGATCCTCTCAG | 840 |
| 841 | Qy | GGACCTGTGTGCTCTGTCTGGAGAGAGAAACGGCCACGCGAGCTGCTTCTTTAAAGCTTA | 900 |
| 841 | Db | GGACCTGTGTGCTCTGTCTGGAGAGAGAAACGGCCACGCGAGCTGCTTCTTTAAAGCTTA | 900 |
| 901 | Qy | ATTCCGAAGAAGCTGCTCTCGAGCTATCTTTGCTCATCAACACGCGCCACCGCTTTG | 960 |
| 901 | Db | ATTCCGAAGAAGCTGCTCTCGAGCTATCTTTGCTCATCAACACGCGCCACCGCTTTG | 960 |
| 961 | Qy | GAGAGCTCCACCGTCTGCTGAGCAACCCAGGCATGTGGAGATCTTGGGCTTCTCTGAG | 1020 |
| 961 | Db | GAGAGCTCCACCGTCTGCTGAGCAACCCAGGCATGTGGAGATCTTGGGCTTCTCTGAG | 1020 |
| 1021 | Qy | GCAGAAAGGAAGAAATCTTCTACAGTATTTCCAAATGCAAGCAGCGCGGCCCAAGTC | 1080 |
| 1021 | Db | GCAGAAAGGAAGAAATCTTCTACAGTATTTCCAAATGCAAGCAGCGCGGCCCAAGTC | 1080 |
| 1081 | Qy | TTCAATTACGTGAGGACAAACGAGCTCTCTTACCATGTGCTTCTGTCCTCTGGTGTC | 1140 |
| 1081 | Db | TTCAATTACGTGAGGACAAACGAGCTCTCTTACCATGTGCTTCTGTCCTCTGGTGTC | 1140 |
| 1141 | Qy | TGGGTGGTGTGTAACCTGCCATCCAGCAGCAGCTGGAGGTGGGGGGTGTGTGAGACAGCG | 1200 |
| 1141 | Db | TGGGTGGTGTGTAACCTGCCATCCAGCAGCAGCTGGAGGTGGGGGGTGTGTGAGACAGCG | 1200 |
| 1201 | Qy | TCAGAGCAACCACTGCAAGTGTACATGCTTACCTCTGATGCTGATGCAACCAAGCCG | 1260 |
| 1201 | Db | TCAGAGCAACCACTGCAAGTGTACATGCTTACCTCTGATGCTGATGCAACCAAGCCG | 1260 |
| 1261 | Qy | GGGGCCCCCGGCTCCAGCCCCCACCAACAGAGAGGGTGTGTCTCTTGGCGGAGAT | 1320 |
| 1261 | Db | GGGGCCCCCGGCTCCAGCCCCCACCAACAGAGAGGGTGTGTCTCTTGGCGGAGAT | 1320 |
| 1321 | Qy | GGGCTCTGGAAATCAGAAAATCTTATTGAGGACAGGACCTCCGGAAGCACCGCCTAGAC | 1380 |
| 1321 | Db | GGGCTCTGGAAATCAGAAAATCTTATTGAGGACAGGACCTCCGGAAGCACCGCCTAGAC | 1380 |
| 1381 | Qy | GGGGAGAGCTCTGCTCTTCTCAACATGAAATCTTCCAGAAAGGACATCAACTGTGAG | 1440 |
| 1381 | Db | GGGGAGAGCTCTGCTCTTCTCAACATGAAATCTTCCAGAAAGGACATCAACTGTGAG | 1440 |
| 1441 | Qy | AGGTACTACAGCTTCAATCCACTTTGAGTTTCCAGGAATCTTTTCGAGCTATGTACTATATC | 1500 |

PS Example; SEQ ID NO 23; 193pp; English.
xx
CC The invention relates to a novel isolated nucleic acid molecule encoding
CC a PAAD-containing polypeptide. The polypeptide of the invention
CC demonstrates cytoskeletal activity and may be useful for preparing a
CC composition for diagnosing or treating diseases associated with the PAAD
CC domain-containing polypeptide, such as cancer, via gene therapy. The
CC current sequence is that of the human PAAD domain-containing protein full
CC -length cDNA of the invention.
xx
SQ Sequence 3108 BP; 682 A; 883 C; 931 G; 612 T; 0 U; 0 Other;

Query Match 100.0%; Score 3108; DB 12; Length 3108;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTACGAACCGCAGGAGGAGCGGCTCTGTGCGCTGTCTCCACTTGGGAAGAACTC 60
DB 1 ATGCTACGAACCGCAGGAGGAGCGGCTCTGTGCGCTGTCTCCACTTGGGAAGAACTC 60

QY 61 GAGGCTGTGGAATGAAGATTCAATTATACCTGGGGACCGGACAGAGCTGGGAGAA 120
DB 61 GAGGCTGTGGAATGAAGATTCAATTATACCTGGGGACCGGACAGAGCTGGGAGAA 120

QY 121 GGCAGATCCCTCGGGAAGCATGAGAGCGCGTCCCTGGAAATGGCCCAAGCTGCTC 180
DB 121 GGCAGATCCCTCGGGAAGCATGAGAGCGCGTCCCTGGAAATGGCCCAAGCTGCTC 180

QY 181 ATCAACCCATCTCGGGCCAGAGGAGGCTGTGAGGTGGCTCTCAAGCACTTTGAGCGGATA 240
DB 181 ATCAACCCATCTCGGGCCAGAGGAGGCTGTGAGGTGGCTCTCAAGCACTTTGAGCGGATA 240

QY 241 AACAGGAAGACTGTGGGAGAGGACAGAGAGGAGCACTGTGAGGATCCCAAGGAA 300
DB 241 AACAGGAAGACTGTGGGAGAGGACAGAGAGGAGCACTGTGAGGATCCCAAGGAA 300

QY 301 ACCTACAGGACTATGTCGAGGAAATTCGGGCTCATGGAAGACCGCAATGGCGGCTA 360
DB 301 ACCTACAGGACTATGTCGAGGAAATTCGGGCTCATGGAAGACCGCAATGGCGGCTA 360

QY 361 GGGGAATGTCAACCTCAGCCACCGGTAACACCGGCTCTGTGTTGAGGAGCACTCA 420
DB 361 GGGGAATGTCAACCTCAGCCACCGGTAACACCGGCTCTGTGTTGAGGAGCACTCA 420

QY 421 AACCCCATGAGGTCAGAGCAGCTTCTGGACACAGCCCGGACACGAGGAGCCGTG 480
DB 421 AACCCCATGAGGTCAGAGCAGCTTCTGGACACAGCCCGGACACGAGGAGCCGTG 480

QY 481 GGACACCAAGGCTAGCCCATCAAGATAGAGACCCCTCTTTGAGCCAGAGGAGCGCCC 540
DB 481 GGACACCAAGGCTAGCCCATCAAGATAGAGACCCCTCTTTGAGCCAGAGGAGCGCCC 540

QY 541 GAGCCACCGCGCACCGTGGTATGACGCGGGGAGGATAGGCAAGTCCATCTGGCA 600
DB 541 GAGCCACCGCGCACCGTGGTATGACGCGGGGAGGATAGGCAAGTCCATCTGGCA 600

QY 601 CACAGGTCATCTGCTGAGTGGCGGACGGAGGCTTCTTCCAGGAGATTTGATTATCTC 660
DB 601 CACAGGTCATCTGCTGAGTGGCGGACGGAGGCTTCTTCCAGGAGATTTGATTATCTC 660

QY 661 TTCTACATCAACTGAGGAGATGAACAGAGTGGCCACGGAATGACATGCAAGCTC 720
DB 661 TTCTACATCAACTGAGGAGATGAACAGAGTGGCCACGGAATGACATGCAAGCTC 720

QY 721 ATCTTACAGTGTGGCTGAGCCGAGCGGCTCTCCAGGAGCTATCCGAGTTCCCGAG 780
DB 721 ATCTTACAGTGTGGCTGAGCCGAGCGGCTCTCCAGGAGCTATCCGAGTTCCCGAG 780

QY 781 GCGCTCTCTTTCATCATCAAGCGGCTTCGATGAGCTCAAGGCTTCTTTCCAGATCCTCAG 840
DB 781 GCGCTCTCTTTCATCATCAAGCGGCTTCGATGAGCTCAAGGCTTCTTTCCAGATCCTCAG 840

QY 841 GGACCCCTGCTGCTGTGGGAGGAGAAACGGGCCACGGAGCTGCTTTTAAAGCTTA 900

DB 841 GGACCCCTGCTGCTGTGGAGGAGAAACGGGCCACGGAGCTGCTTTTAAAGCTTA 900

QY 901 ATTGGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACAGGCCCCACGGCTTTG 960
DB 901 ATTGGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACAGGCCCCACGGCTTTG 960

QY 961 GAGAGCTCCACCCGCTGCTGAGGACACCCAGGATGTGGAGATCCTGGGCTTCTCTGAG 1020
DB 961 GAGAGCTCCACCCGCTGCTGAGGACACCCAGGATGTGGAGATCCTGGGCTTCTCTGAG 1020

QY 1021 GCAGAAAGGAAGGAATATTCTTACAAGATTTCACAATGACAGAGCAGCGGGCCAAAGTC 1080
DB 1021 GCAGAAAGGAAGGAATATTCTTACAAGATTTCACAATGACAGAGCAGCGGGCCAAAGTC 1080

QY 1081 TTCAATTAAGTACGAGGAGCAACGAGGCTCTCTTACCATGTCTTCGTCCTCCCTGCTGTC 1140
DB 1081 TTCAATTAAGTACGAGGAGCAACGAGGCTCTCTTACCATGTCTTCGTCCTCCCTGCTGTC 1140

QY 1141 TGGGTGGTGTGTACCTGCTCTCCAGCAGCAGCTGGAGGTGGGGCTGTGAGACAGACG 1200
DB 1141 TGGGTGGTGTGTACCTGCTCTCCAGCAGCAGCTGGAGGTGGGGCTGTGAGACAGACG 1200

QY 1201 TCCAGGAACACCACTGCGAGTGTACATGTCTTACCTGCTGAGTCTGATGCAACCCAAAGCG 1260
DB 1201 TCCAGGAACCACTGCGAGTGTACATGTCTTACCTGCTGAGTCTGATGCAACCCAAAGCG 1260

QY 1261 GGGGCCCGCGCTCCAGCCCCACCAACAGAGAGGCTGTGCTCTTGGCGGACAGAT 1320
DB 1261 GGGGCCCGCGCTCCAGCCCCACCAACAGAGAGGCTGTGCTCTTGGCGGACAGAT 1320

QY 1321 GGGCTCTGGAATCAGAAATCCTATTGAGGAGCAGGACCTCCGGAAGCAGCGGCTAGAC 1380
DB 1321 GGGCTCTGGAATCAGAAATCCTATTGAGGAGCAGGACCTCCGGAAGCAGCGGCTAGAC 1380

QY 1381 GGGGAAGAGCTCTGCTGCTTCCATGACATGACATCTTCCAGAGACATCAACTGTGAG 1440
DB 1381 GGGGAAGAGCTCTGCTGCTTCCATGACATGACATCTTCCAGAGACATCAACTGTGAG 1440

QY 1441 AGTACTACAGCTTCACTTCCACTTGGATTTCCAGGAATTTCTTGGAGCTATGTAATATC 1500
DB 1441 AGTACTACAGCTTCACTTCCACTTGGATTTCCAGGAATTTCTTGGAGCTATGTAATATC 1500

QY 1501 CTGGAACGAGGGAGGCGGGGAGCCAGACGAGGAGCTGACAGGCTGTGACCCGAG 1560
DB 1501 CTGGAACGAGGGAGGCGGGGAGCCAGACGAGGAGCTGACAGGCTGTGACCCGAG 1560

QY 1561 TAGCGCTTTTCTGAAAGGAGCTTCTGCGCACTCACCAGCGCTTCTGTTGGACTCTCTG 1620
DB 1561 TAGCGCTTTTCTGAAAGGAGCTTCTGCGCACTCACCAGCGCTTCTGTTGGACTCTCTG 1620

QY 1621 AACGAGGAGACCAAGGAGCCCTTGGAGAGAGTCTCTGTGGAAGGTCTCGCGGACATC 1680
DB 1621 AACGAGGAGACCAAGGAGCCCTTGGAGAGAGTCTCTGTGGAAGGTCTCGCGGACATC 1680

QY 1681 AAGATGGAACCTGTTGAGTGGATCCAAAGCAAGCTCAGAGCGGCTCCACCTGCGAG 1740
DB 1681 AAGATGGAACCTGTTGAGTGGATCCAAAGCAAGCTCAGAGCGGCTCCACCTGCGAG 1740

QY 1741 CAGGGCTCTTGGAGTTCCTGAGTGTGACGATCCAGAGATCCAGGAGGAGGATTTATCCAG 1800
DB 1741 CAGGGCTCTTGGAGTTCCTGAGTGTGACGATCCAGAGATCCAGGAGGAGGATTTATCCAG 1800

QY 1801 CAGGCCCTGAGCCACTTCCAGGTGATCGTGGTCAAGCAACTTCCCTCCAGATGGAGCAC 1860
DB 1801 CAGGCCCTGAGCCACTTCCAGGTGATCGTGGTCAAGCAACTTCCCTCCAGATGGAGCAC 1860

QY 1861 ATGCTCTCTGCTGCTGAGGAGCGCTGAGGAGCGCCAGGCTGTGCACTTGTATGGC 1920
DB 1861 ATGCTCTCTGCTGCTGAGGAGCGCTGAGGAGCGCCAGGCTGTGCACTTGTATGGC 1920

QY 1921 GCCACCTTACAGCGCGGAGGAGAGACCGCGAGGCTGCTCCGAGAGGAGCGACACGCTG 1980

Db 1921 GCCACCTACAGCGCGGACGCGGGAAGACCGCGGAGGTGCTCCGAGGAGCGCACACGCTG 1980
Qy 1981 TTGTGTCAGCTCAGACACAGAGAGACCGTCTTGTGTGACGCTTACAGTGAACATCTGGCA 2040
Db 1981 TTGTGTCAGCTCAGACACAGAGAGACCGTCTTGTGTGACGCTTACAGTGAACATCTGGCA 2040
Qy 2041 CGGGCCCTGTGCACCAATCCAAACCTGTATAGAGTGTCTGTATCCGAAATGCCCTGGGC 2100
Db 2041 CGGGCCCTGTGCACCAATCCAAACCTGTATAGAGTGTCTGTATCCGAAATGCCCTGGGC 2100
Qy 2101 AGCCGGGGGGTGAAGTCTCTGTCAAGGACTCAGACACCCCAACTGCARAATTCAGAAC 2160
Db 2101 AGCCGGGGGGTGAAGTCTCTGTCAAGGACTCAGACACCCCAACTGCARAATTCAGAAC 2160
Qy 2161 CTGAGGCTGAAGAGGTGCGGCATCTCCAGCTCAGCTCGGAGGACCTCTCTGCACTCTC 2220
Db 2161 CTGAGGCTGAAGAGGTGCGGCATCTCCAGCTCAGCTCGGAGGACCTCTCTGCACTCTC 2220
Qy 2221 ATAGCCAAATGAATTTGACAAGATGATCTCAGTGGCAACGGGGTTGGATTTCCAGGC 2280
Db 2221 ATAGCCAAATGAATTTGACAAGATGATCTCAGTGGCAACGGGGTTGGATTTCCAGGC 2280
Qy 2281 ATGATGCTGCTTTCGGAGGCTCGGCATCCAGTCCAGTCCAGTCCAGATGATTCAGTTG 2340
Db 2281 ATGATGCTGCTTTCGGAGGCTCGGCATCCAGTCCAGTCCAGTCCAGATGATTCAGTTG 2340
Qy 2341 AGGAAGTGTCACTGAGTCCGGGCTTGTGAGAGATGGCTTCTGTGCTCGGACCAAC 2400
Db 2341 AGGAAGTGTCACTGAGTCCGGGCTTGTGAGAGATGGCTTCTGTGCTCGGACCAAC 2400
Qy 2401 CCACATCTGTTGATTTGGAACCTGACAGGAATGCATGAGGATTTGGGCTCAGGTTA 2460
Db 2401 CCACATCTGTTGATTTGGAACCTGACAGGAATGCATGAGGATTTGGGCTCAGGTTA 2460
Qy 2461 CTATGCCAGGACTGAGCACCAGTCTGACAGTACGAGTACGAGTCTGAGTCTGAGTCTGC 2520
Db 2461 CTATGCCAGGACTGAGCACCAGTCTGACAGTACGAGTACGAGTCTGAGTCTGAGTCTGC 2520
Qy 2521 CGCTCTACTGCTGCTGTGACGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 2580
Db 2521 CGCTCTACTGCTGCTGTGACGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 2580
Qy 2581 AGAGAGTGGACCTGAGCTGAATGAGTGGGGACCTTCGGGCTGCTGCTGTGTGAG 2640
Db 2581 AGAGAGTGGACCTGAGCTGAATGAGTGGGGACCTTCGGGCTGCTGCTGTGTGAG 2640
Qy 2641 GGCCTCAGGATCCACACCTCAAGCTCCAGACCTTCGGGCTGCGGCTGCGGCTGGC 2700
Db 2641 GGCCTCAGGATCCACACCTCAAGCTCCAGACCTTCGGGCTGCGGCTGCGGCTGGC 2700
Qy 2701 TCTGCCGCTGTGAGGCTTCTTCTGTGTGTGCTCCAGGCAACCAACCTCCGGGAGCTG 2760
Db 2701 TCTGCCGCTGTGAGGCTTCTTCTGTGTGTGCTCCAGGCAACCAACCTCCGGGAGCTG 2760
Qy 2761 GACTTGAATTTCAACGACCTGGGAGACTGGGCTGTGGTGTGCTGCTGAGGGCTGCAA 2820
Db 2761 GACTTGAATTTCAACGACCTGGGAGACTGGGCTGTGGTGTGCTGCTGAGGGCTGCAA 2820
Qy 2821 CATCCGCTCAGACTCCAGAACTGTGGCTGTGATAGTCTGGGCTCAGACCAAGCT 2880
Db 2821 CATCCGCTCAGACTCCAGAACTGTGGCTGTGATAGTCTGGGCTCAGACCAAGCT 2880
Qy 2881 TGTGAGATCTTTACTTCCCTGGGATCAACAGACCTTTGACGACCTTTTACTGACC 2940
Db 2881 TGTGAGATCTTTACTTCCCTGGGATCAACAGACCTTTGACGACCTTTTACTGACC 2940
Qy 2941 AACACGCTTAGGGGACACAGGTGTCCGACTGCTTTTGAAGCGGCTGAGCCATCTGGC 3000
Db 2941 AACACGCTTAGGGGACACAGGTGTCCGACTGCTTTTGAAGCGGCTGAGCCATCTGGC 3000
Qy 3001 TGCAAACTCCAGTCTCTGTTTATTTGGGATGACCTGGAATAAATGACCCACAGTAGG 3060
Db 3001 TGCAAACTCCAGTCTCTGTTTATTTGGGATGACCTGGAATAAATGACCCACAGTAGG 3060

Qy 3061 TTGSCAGCGCTTCGAGTAACAAAACCTTATTTTGGACATTGGCTGCTGA 3108
Db 3061 TTGSCAGCGCTTCGAGTAACAAAACCTTATTTTGGACATTGGCTGCTGA 3108

RESULT 3

AAAL47129 standard; DNA; 3300 BP.

XX AC AAL47129;

XX XX 20-AUG-2002 (first entry)

DE DE Pyrin domain containing protein NALP3/PY5-hs coding sequence.

KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW nephrotropic; osteopathic; nontropic; intracellular signal transduction;
KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW osteoarthritis; glomerulonephritis; gene; ds.

XX OS Unidentified.

XX XX WO200240668-A2.

XX XX 23-MAY-2002.

XX XX 30-OCT-2001; 2001WO-EP012545.

XX XX 15-NOV-2000; 2000DE-01056687.

XX XX 30-NOV-2000; 2000DE-01059595.

XX XX (APOT-) APOTEC RES & DEV LTD.

XX XX Tschopp J, Marton F;

XX XX WPI; 2002-427093/45.

XX XX P-PSDB; AAO17857.

XX XX New DNA encoding protein with pyrin domain, useful for treating diseases involving impaired signal transduction, particularly inflammation, also proteins and antibodies.

XX XX Claim 5; Fig 1; 116pp; German.

XX XX The present invention relates the DNA and their encoded proteins, where the proteins contain at least one PYD (pyrin) domain. These can be used to treat diseases associated with impaired intracellular signal transduction, particularly inflammation such as psoriasis, arteriosclerosis, bacterial or viral infections (particularly meningitis and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma, sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's and Parkinson's diseases. The present sequence is a coding sequence of the invention

XX XX Sequence 3300 BP; 726 A; 943 C; 979 G; 652 T; 0 U; 0 Other;

Query Match 99.4%; Score 3057; DB 6; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3057; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTAGAACCCGAGCAGGACGGCTCTGTGCTGCTCCACTTACTTGGAGAACTC 60
Db 1 ATGCTAGAACCCGAGCAGGACGGCTCTGTGCTGCTCCACTTACTTGGAGAACTC 60

Qy 61 GAGGCTCTGGAACCTGAAGAAGTTCAAGTATATCTGGGACCGGCACAGAGCTGGGAA 120
Db 61 GAGGCTCTGGAACCTGAAGAAGTTCAAGTATATCTGGGACCGGCACAGAGCTGGGAA 120

Qy 121 GGCAAGATCCCTGGGGGAAGCATGGAGAAGCCGGTCCCTCGGAAATGGCCCACTGCTC 180

Db 121 GGAAGATCCCTGGGGAAGCATGGGAAGCCGGTCCCTGGAAATGGCCAGCTGCTC 180
Qy 181 ATCAACCCACATTCGGGCCAGAGAGGCTGGAGGTGGCTCTCAGACACCTTTGAGCGGATA 240
Db 181 ATCAACCCACATTCGGGCCAGAGAGGCTGGAGGTGGCTCTCAGACACCTTTGAGCGGATA 240
Qy 241 AACAGGAAGGACCTGTGGGAGAGAGGACAGAGAGGACCTGTGTGAGGGATCCCHAGGAA 300
Db 241 AACAGGAAGGACCTGTGGGAGAGAGGACAGAGAGGACCTGTGTGAGGGATCCCHAGGAA 300
Qy 301 ACCTACAGGACATGTCCGAGGAATTCGGCTCATGGAACACCGCAATGCGCCCTTA 360
Db 301 ACCTACAGGACATGTCCGAGGAATTCGGCTCATGGAACACCGCAATGCGCCCTTA 360
Qy 361 GGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTCCTGTGTGAGAGGACACTCA 420
Db 361 GGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTCCTGTGTGAGAGGACACTCA 420
Qy 421 AACCCCATGCAAGGTCCAGAGACAGCTTCTGGACACAGGCCGGGACACGCGAGGACCGTG 480
Db 421 AACCCCATGCAAGGTCCAGAGACAGCTTCTGGACACAGGCCGGGACACGCGAGGACCGTG 480
Qy 481 GGACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGACGAGGAGCGCCCC 540
Db 481 GGACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGACGAGGAGCGCCCC 540
Qy 541 GAGCCACCGCACCGTGGTCAATCAAGCGCGCGAGGGATAGCAAGTCCATGTGGCA 600
Db 541 GAGCCACCGCACCGTGGTCAATCAAGCGCGCGAGGGATAGCAAGTCCATGTGGCA 600
Qy 601 CACAAGGTGATGCTGGACTGGCGGAGCGGAGCTCTTCCAGGACGATTTGATTAATCTC 660
Db 601 CACAAGGTGATGCTGGACTGGCGGAGCGGAGCTCTTCCAGGACGATTTGATTAATCTC 660
Qy 661 TTCTACATCAACTGCAGGGAGATGAACAGAGTGCACCGAATGCAGCATGCAGACCTC 720
Db 661 TTCTACATCAACTGCAGGGAGATGAACAGAGTGCACCGAATGCAGCATGCAGACCTC 720
Qy 721 ATCTTCAGCTGTGCTGAGCCAGCGCGCTCTCCAGGAGCTCATCCGAGTTCGCGAG 780
Db 721 ATCTTCAGCTGTGCTGAGCCAGCGCGCTCTCCAGGAGCTCATCCGAGTTCGCGAG 780
Qy 781 GGCCTCTTTTCATCATCGAGGCTTCGATGAGCTCAAGCTCTTTCCAGATCTCAG 840
Db 781 GGCCTCTTTTCATCATCGAGGCTTCGATGAGCTCAAGCTCTTTCCAGATCTCAG 840
Qy 841 GGACCTGTGCTCTGTGGGAGGAGAAACGGCCACAGGAGCTGCTTTTAAAGCTTA 900
Db 841 GGACCTGTGCTCTGTGGGAGGAGAAACGGCCACAGGAGCTGCTTTTAAAGCTTA 900
Qy 901 ATTGGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACAGGCCACCGCTTG 960
Db 901 ATTGGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACAGGCCACCGCTTG 960
Qy 961 GAGAGCTCCACCGTCTGCTGAGCACCCAGGAGTGGAGATCCTGGGCTTCTCTGAG 1020
Db 961 GAGAGCTCCACCGTCTGCTGAGCACCCAGGAGTGGAGATCCTGGGCTTCTCTGAG 1020
Qy 1021 GCAGAAAGGAAGGAATACTTCTAAGATATTTCCAAJTGACAGACAGCGGGCCAAAGTC 1080
Db 1021 GCAGAAAGGAAGGAATACTTCTAAGATATTTCCAAJTGACAGACAGCGGGCCAAAGTC 1080
Qy 1081 TTCAATTACGTGAGGGACAAAGAGCTCTTTCAACATGTGCTGCTCCCTGCTGTGTC 1140
Db 1081 TTCAATTACGTGAGGGACAAAGAGCTCTTTCAACATGTGCTGCTCCCTGCTGTGTC 1140
Qy 1141 TGGGTGTGTGTACCTGCTCCAGACAGAGCTGGAGGCTGGGGGCTGTGTGAGACAGAG 1200
Db 1141 TGGGTGTGTGTGTACCTGCTCCAGACAGAGCTGGAGGCTGGGGGCTGTGTGAGACAGAG 1200
Qy 1201 TCCAGACACCACTGAGTGTACATGCTCTACCTGTGAGTCTGTATGCAACCAAGCCG 1260

Db 1201 TCCAGACACCACTGAGTGTACATGCTCTACCTGTGAGTCTGTATGCAACCAAGCCG 1260
Qy 1261 GGGGCCCGCGCTTCAGCCCCACCCAAACAGAGAGGTTGTCTCTCTTGGCGGAGAT 1320
Db 1261 GGGGCCCGCGCTTCAGCCCCACCCAAACAGAGAGGTTGTCTCTTCTTGGCGGAGAT 1320
Qy 1321 GGGCTCTGGAATCAGAAATCTTATTTGAGGAGGAGGACCTTCGGAAGACAGCGGCTAGAC 1380
Db 1321 GGGCTCTGGAATCAGAAATCTTATTTGAGGAGGAGGACCTTCGGAAGACAGCGGCTAGAC 1380
Qy 1381 GGGGAAGAGCTCTTCGCTTCTTAAATGAAATCTTCCAGAAAGACATCAACTGTGAG 1440
Db 1381 GGGGAAGAGCTCTTCGCTTCTTAAATGAAATCTTCCAGAAAGACATCAACTGTGAG 1440
Qy 1441 AGGTACTACAGCTTACATCTGAGTTTCCAGGAATCTTTGAGGCTATGTACTATATC 1500
Db 1441 AGGTACTACAGCTTACATCTGAGTTTCCAGGAATCTTTGAGGCTATGTACTATATC 1500
Qy 1501 CTGGAAGAGGAGGCGGGGCGAGCCAGACAGGACGTGACAGGCTGTGAGCCGAG 1560
Db 1501 CTGGAAGAGGAGGCGGGGCGAGCCAGACAGGACGTGACAGGCTGTGAGCCGAG 1560
Qy 1561 TACGGGTTTTCTGAAAGAGCTTCTGAGCTCCTGAGCTCCTGTTGGACTCTCTG 1620
Db 1561 TACGGGTTTTCTGAAAGAGCTTCTGAGCTCCTGAGCTCCTGTTGGACTCTCTG 1620
Qy 1621 AACGAGGAGACGAGGAGCCTGAGAGAGTCTCTGCTGGAAGGCTCTCGCGGACATC 1680
Db 1621 AACGAGGAGACGAGGAGCCTGAGAGAGTCTCTGCTGGAAGGCTCTCGCGGACATC 1680
Qy 1681 AAGATGGAACCTGTTGAGTGGATCCAAAGCAAGCTCAGAGGACGGCTCCACCTTGAG 1740
Db 1681 AAGATGGAACCTGTTGAGTGGATCCAAAGCAAGCTCAGAGGACGGCTCCACCTTGAG 1740
Qy 1741 CAGGGCTCTTGGAGCTTCTTACAGTCTTGTACAGATCCAGAGAGGAGTGTATCCAG 1800
Db 1741 CAGGGCTCTTGGAGCTTCTTACAGTCTTGTACAGATCCAGAGAGGAGTGTATCCAG 1800
Qy 1801 CAGGGCTCTGAGCCTTCCAGGCTGATCGTGGTCAAGCAATTCCTCCAAAGTGGAGCAC 1860
Db 1801 CAGGGCTCTGAGCCTTCCAGGCTGATCGTGGTCAAGCAATTCCTCCAAAGTGGAGCAC 1860
Qy 1861 ATGGTCTCTGTTGAAAGCTGAGAGGCGCCAGGTCGCTGACCTGTATGTC 1920
Db 1861 ATGGTCTCTGTTGAAAGCTGAGAGGCGCCAGGTCGCTGACCTGTATGTC 1920
Qy 1921 GGCACCTACAGCGCGGAGGAGACCGCGGAGGTGCTCCGAGAGCGGACACGCTG 1980
Db 1921 GGCACCTACAGCGCGGAGGAGACCGCGGAGGTGCTCCGAGAGCGGACACGCTG 1980
Qy 1981 TTGGTGCAGCTCAGACAGAGGACCGTTCGTGGAAGCGCTTACAGTGAATCTGGCA 2040
Db 1981 TTGGTGCAGCTCAGACAGAGGACCGTTCGTGGAAGCGCTTACAGTGAATCTGGCA 2040
Qy 2041 GCGGCCCTGTGACCAATCCAAACCTGATAGAGTCTCTGTACGAAATGCCCTGGGC 2100
Db 2041 GCGGCCCTGTGACCAATCCAAACCTGATAGAGTCTCTGTACGAAATGCCCTGGGC 2100
Qy 2101 AGCGGGGGGTGAAGTGTCTCTCAAGGACTCAGACACCCCACTGCAAACTCAGAAC 2160
Db 2101 AGCGGGGGGTGAAGTGTCTCTCAAGGACTCAGACACCCCACTGCAAACTCAGAAC 2160
Qy 2161 CTGAGGCTGAAGAGGTGCGCATCTCAGCTCAGCTGCGAGGACCTCTCTGAGCTCTC 2220
Db 2161 CTGAGGCTGAAGAGGTGCGCATCTCAGCTCAGCTGCGAGGACCTCTCTGAGCTCTC 2220
Qy 2221 ATAGCCCAATGAAGATTTGCAAGGATCTCAGTGGGAAACGGCGTGTGATTCAGAGC 2280
Db 2221 ATAGCCCAATGAAGATTTGCAAGGATCTCAGTGGGAAACGGCGTGTGATTCAGAGC 2280
Qy 2281 ATGATGCTGCTTTGCGAGGCGCTCGGCATCTCCAGTGCAGGCTGAGATGATTCAGTTG 2340
Db 2281 ATGATGCTGCTTTGCGAGGCGCTCGGCATCTCCAGTGCAGGCTGAGATGATTCAGTTG 2340

```
QY 2341 AGGAAGTGTCAAGTCGAGTCGGGGCTTTGTCAAGAGATGGCTTTGTGCTCGGCACCAAC 2400
Db |||||
QY 2341 AGGAAGTGTCAAGTCGAGTCGGGGCTTTGTCAAGAGATGGCTTTGTGCTCGGCACCAAC 2400
Db |||||
QY 2401 CCACATCTGTTGAGTTGGACTGACAGAAATGACAGGAGATTTGGGCTGAGGTTA 2460
Db |||||
QY 2401 CCACATCTGTTGAGTTGGACTGACAGAAATGACAGGAGATTTGGGCTGAGGTTA 2460
Db |||||
QY 2461 CTATGCCAGGACTGAGGACCCAGTCTGCGAGACTACCGACTTTGTGGCTGAAGATCGC 2520
Db |||||
QY 2461 CTATGCCAGGACTGAGGACCCAGTCTGCGAGACTACCGACTTTGTGGCTGAAGATCGC 2520
Db |||||
QY 2521 GGCCTCACTGTGCTGCTGAGCGAGTGCCTCAACTCTCAGTGTGAACAGAGCGTG 2580
Db |||||
QY 2521 GGCCTCACTGTGCTGCTGAGCGAGTGCCTCAACTCTCAGTGTGAACAGAGCGTG 2580
Db |||||
QY 2581 AGAGAGCTGAGCTGAGCTGANTGAGCTGGGGACCTCGGGTCTGCTGTGTGAG 2640
Db |||||
QY 2581 AGAGAGCTGAGCTGAGCTGANTGAGCTGGGGACCTCGGGTCTGCTGTGTGAG 2640
Db |||||
QY 2641 GGCCTCAGGATCCACAGTCAAGCTCCAGACCCTGCGGTTGGGCATCTGCCGCTGGGC 2700
Db |||||
QY 2641 GGCCTCAGGATCCACAGTCAAGCTCCAGACCCTGCGGTTGGGCATCTGCCGCTGGGC 2700
Db |||||
QY 2701 TCTGCCGCTGTGAGGGTCTTTCTGTGTGCTCCAGGCCAACCAACCTCCGGAGCTG 2760
Db |||||
QY 2701 TCTGCCGCTGTGAGGGTCTTTCTGTGTGCTCCAGGCCAACCAACCTCCGGAGCTG 2760
Db |||||
QY 2761 GACTTGAGTTTCAACGACTGGGAGACTGGGGCTGTGGTGTGGCTGAGGGCTGCAA 2820
Db |||||
QY 2761 GACTTGAGTTTCAACGACTGGGAGACTGGGGCTGTGGTGTGGCTGAGGGCTGCAA 2820
Db |||||
QY 2821 CATCCCGCTGCAGACTCCAGAACTGTGGTGTGATAGTGTGGCTCACAAGCAAGCT 2880
Db |||||
QY 2821 CATCCCGCTGCAGACTCCAGAACTGTGGTGTGATAGTGTGGCTCACAAGCAAGCT 2880
Db |||||
QY 2881 TGTGAGAACTTTACTTCACTCCCTGGGATCAACAGACTTTGACCGACTTTTACCTGACC 2940
Db |||||
QY 2881 TGTGAGAACTTTACTTCACTCCCTGGGATCAACAGACTTTGACCGACTTTTACCTGACC 2940
Db |||||
QY 2941 AACACGCCCTAGGGACACAGTGTCCGACTGCTTTTCAAGCGGCTGAGCATCTCTGGC 3000
Db |||||
QY 2941 AACACGCCCTAGGGACACAGTGTCCGACTGCTTTTCAAGCGGCTGAGCATCTCTGGC 3000
Db |||||
QY 3001 TGCAAACTCCGAGTCTCTGTTATTGCGGATGAGACCTGAAATAAAATGACCCACAGT 3057
Db |||||
QY 3001 TGCAAACTCCGAGTCTCTGTTATTGCGGATGAGACCTGAAATAAAATGACCCACAGT 3057
Db |||||

RESULT 4
ADP47723
ID ADP47723 standard; cDNA; 3731 BP.
XX
AC ADP47723;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human Monarch-1 (CATERPILLER 19.3) full-length cDNA.
XX
KW Monarch-1; CATERPILLER 11.2; caspase recruitment domain;
KW CARD, transcription enhancer, R-binding, pyrin, lots of leucine repeat;
KW CATERPILLER 11.3; CATERPILLER 16.1; CATERPILLER 16.2; CIAS1;
KW cold-induced autoinflammatory syndrome 1; antiinflammatory; cytostatic;
KW inflammatory disease; cancer; gene therapy; human; gene; ss; purine;
KW CATERPILLER 19.3; chromosome 19q13; multiple sclerosis.
XX
OS Homo sapiens.
XX
PN WO2004034093-A2.
XX
PD 22-APR-2004.
XX
```

```
PF 30-APR-2003; 2003MO-US013562.
PR 30-APR-2002; 2002US-0376626P.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Ting JY, Linhoff MW, Harton JA, Williams KL, Lich J, O'Connor W;
PI Moore CB, Davis B, Brickey J, Conti B, Zhang J, Zhu X;
XX
DR WPI; 2004-348215/32.
DR P-PSDB; ADP47724.
XX
PT New nucleic acid encoding Monarch-1, CATERPILLER 11.2, CATERPILLER 11.3,
PT CATERPILLER 16.2, CATERPILLER 16.2 or CIAS1 polypeptide, useful in
PT preparing a composition for treating inflammatory disease or cancer.
XX
PS Claim 4; SEQ ID NO 1; 205pp; English.
XX
CC The invention relates to a novel isolated nucleic acid encoding a Monarch
CC -1, CATERPILLER (CARD [caspase recruitment domain], transcription
CC enhancer, R (purine)-binding, pyrin, lots of leucine repeats) 11.2,
CC CATERPILLER 11.3, CATERPILLER 16.1, CATERPILLER 16.2 or CIAS1 [cold-
CC induced autoinflammatory syndrome 1] polypeptide comprising the amino
CC acids encoded by exon 6 and lacking the amino acids encoded by exon 4 or
CC its fragment. The nucleic acid of the invention demonstrates
CC antiinflammatory and cytostatic activities and may be useful in preparing
CC a composition for treating an inflammatory disease or cancer, possibly
CC via gene therapy. The current sequence is that of the human Monarch-1
CC (CATERPILLER 19.3) full-length cDNA of the invention which is located on
CC chromosome 19q13, in the multiple sclerosis susceptibility region.
XX
SQ Sequence 3731 BP; 842 A; 1046 C; 1091 G; 752 T; 0 U; 0 Other;

Query Match 97.1%; Score 3017; DB 12; Length 3731;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 3108; Conservative 0; Mismatches 0; Indels 81; Gaps 1;

QY 1 ATGCTACGAACCCGACGAGGACGGCTCTGTGCTGTCCACTTCTTGGAGAACTC 60
Db 221 ATGCTACGAACCCGACGAGGACGGCTCTGTGCTGTCCACTTCTTGGAGAACTC 280

QY 61 GAGCTGTGNACTGAAGATTCAAGTTATACCTGGGGACCGGACAGAGCTGGGAA 120
Db 281 GAGCTGTGNACTGAAGATTCAAGTTATACCTGGGGACCGGACAGAGCTGGGAA 340

QY 121 GGCAAGATCCCTCGGGAAGCATCGAAGCCGCTCCCTGGAAATGCCCCAGCTGTC 180
Db 341 GGCAAGATCCCTCGGGAAGCATCGAAGCCGCTCCCTGGAAATGCCCCAGCTGTC 400

QY 181 ATCAACCCACTTCGGGCCAGAGAGGCGCTTGGAGTTGGCTCTCAGCACCTTTGAGCGGATA 240
Db 401 ATCAACCCACTTCGGGCCAGAGAGGCGCTTGGAGTTGGCTCTCAGCACCTTTGAGCGGATA 460

QY 241 AACAGGAAGACCTTGGGAGAGAGGACAGAGAGGACCTTGGTGAGG----- 288
Db 461 AACAGGAAGACCTTGGGAGAGAGGACAGAGAGGACCTTGGTGAGGAGATACCCACCT 520

QY 289 ----- 288
Db 521 GGTGGCCGCTCTCACTTGGGAACCACTCAACATGCCCTTCTGGAAGTCTCTCTTGTCACT 580

QY 289 -----GATCCCCAGGAACCTACAGGAGTATGTCCGAGGAAATTCGGGTCTATG 339
Db 581 CCAAGAAAAGATCCCCAGGAACCTACAGGAGTATGTCCGAGGAAATTCGGGTCTATG 640

QY 340 GAAGACCGCAATGGCGGCTTAGGGAAATGTGTCAACCTCAGCACCGGTACACCCGGTCTC 399
Db 641 GAAGACCGCAATGGCGGCTTAGGGAAATGTGTCAACCTCAGCACCGGTACACCCGGTCTC 700

QY 400 CTGCTGTGTGAGGAGACACTCAAAACCCATGAGGTCCAGAGCAGAGCTTCTGAGACACAGGC 459
Db 701 CTGCTGTGTGAGGAGACACTCAAAACCCATGAGGTCCAGAGCAGAGCTTCTGAGACACAGGC 760
```

[illegible]

```
Db 2921 GGGGTGCTGCTGTGTGAGGGGCTCAGGCATCCACGTGGAAGCTCCAGACCTCGCG 2980
Qy 2680 TTGGGCATCTGCGGCTGGGCTCTGCGCCCTGTGTGAGGCTCTTTCTGTGTGCTCAGGCC 2739
Db 2981 TTGGGCATCTGCGGCTGGGCTCTGCGCCCTGTGTGAGGCTCTTTCTGTGTGCTCAGGCC 3040
Qy 2740 AACCAAACTTCCGGAGCTGAGCTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGG 2799
Db 3041 AACCAAACTTCCGGAGCTGAGCTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGG 3100
Qy 2800 TTGCTGGCTGAGGGCTCAACATCCCGCTGCGAGCTCCAGAACTGTGGCTGGATGC 2859
Db 3101 TTGCTGGCTGAGGGCTCAACATCCCGCTGCGAGCTCCAGAACTGTGGCTGGATGC 3160
Qy 2860 TGTGGCTTACAGCAAGCTTGTGAGAACTTTTACTTTCACCTCGGGATCAACAGACC 2919
Db 3161 TGTGGCTTACAGCAAGCTTGTGAGAACTTTTACTTTCACCTCGGGATCAACAGACC 3220
Qy 2920 TTGACCGACCTTACCTGACCAACACCGCCCTAGGGGACAGGTGTCGACTGCTTTC 2979
Db 3221 TTGACCGACCTTACCTGACCAACACCGCCCTAGGGGACAGGTGTCGACTGCTTTC 3280
Qy 2980 AAGCGGCTGAGCCATCTCGGCTGCAAACTCCGAGTCTCTGTTATTTGGATGACCTG 3039
Db 3281 AAGCGGCTGAGCCATCTCGGCTGCAAACTCCGAGTCTCTGTTATTTGGATGACCTG 3340
Qy 3040 AATAAATGACCCACAGTAGGTGTGGCAGCGCTTCGAGTAACAAAACCTTATTTGGACAT 3099
Db 3341 AATAAATGACCCACAGTAGGTGTGGCAGCGCTTCGAGTAACAAAACCTTATTTGGACAT 3400
Qy 3100 GGCTGTGA 3108
Db 3401 GGCTGTGA 3409

RESULT 5
AAL44363
ID AAL44363 standard; cDNA; 3186 BP.
XX
AC AAL44363;
XX
DT 31-OCT-2002 (first entry)
XX
DE Human PYRIN-8 cDNA sequence #2.
XX
KW Human; gene; ss; gene therapy; PYRIN; stress-related response;
KW apoptotic response; inflammatory response; inflammatory disorder;
KW immune system disorder; Crohn's disease; multiple sclerosis; cancer;
KW leukaemia; autoimmune disorder; arthritis; neurological disease;
KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
KW transcription profiling; PYRIN-8.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3186
FT /tag= a
FT /product= "Human PYRIN-8 protein #2"
XX
FN WO200261049-A2.
XX
PD 08-AUG-2002.
XX
PF 31-JAN-2002; 2002WO-US002967.
XX
PR 31-JAN-2001; 2001US-0265231P.
PR 10-SEP-2001; 2001US-0318645P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
PA (AMHP) WYETH.
XX
PI Bertin J, Wang W, Blatcher M;
```

XX

WPI; 2002-627477/67.

P-PSDB; AAO15590.

XX

New PYRIN polypeptides and nucleic acids useful for modulating and diagnosing stress-related, apoptotic and inflammatory responses, or for treating inflammatory and immune system disorders, cancers, or neurological diseases.

XX

Claim 4; Fig 8; 167pp; English.

PS

The invention comprises the amino acid and coding sequences of human PYRIN proteins. The PYRIN protein and DNA sequences of the invention are useful for modulating and diagnosing stress-related, apoptotic and inflammatory responses. The PYRIN protein and DNA sequences are useful for treating inflammatory disorders and immune system disorders (e.g. Crohn's disease, reactive arthritis, multiple sclerosis, contact dermatitis, psoriasis, graft rejection, allergies, viral infections and bacterial infections); cancer (e.g. leukaemia); autoimmune disorders (e.g. systemic lupus erythematosus and arthritis); and neurological diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN protein and DNA sequences may also be used in screening assays, detection assays (e.g. chromosomal mapping, tissue typing or forensic biology), predictive medicine (e.g. diagnostic assays, clinical trials and pharmacogenomics) and transcription profiling. The present DNA sequence encodes a human PYRIN-8 protein

XX

Sequence 3186 BP; 701 A; 908 C; 945 G; 632 T; 0 U; 0 Other;

SQ

Query Match 96.5%; Score 2997.8; DB 6; Length 3186;

Best Local Similarity 97.3%; Pred. No. 0;

Matches 3103; Conservative 0; Mismatches 2; Indels 84; Gaps 2;

Qy

1 ATGCTACGAAACCCGAGGAGCGGCTCTGTGCGCTGTCCACTTCTGGAGAACTC 60

Db

1 ATGCTACGAAACCCGAGGAGCGGCTCTGTGCGCTGTCCACTTCTGGAGAACTC 60

Qy

61 GAGGCTGTGGAACCTGAAAGAGTTCAAGTTATACCTGGGGACCGGACAGAGCTGGGAA 120

Db

61 GAGGCTGTGGAACCTGAAAGAGTTCAAGTTATACCTGGGGACCGGACAGAGCTGGGAA 120

Qy

121 GCGAAGATCCCTGGGGAAGCATGGAGAGCGGCTCCCTTGGAAATGGCCAGCTGCTC 180

Db

121 GCGAAGATCCCTGGGGAAGCATGGAGAGCGGCTCCCTTGGAAATGGCCAGCTGCTC 180

Qy

181 ATACCCACATTCGGGCCAGAGAGGCTTGGAGTTGGCTCTCAGCAGCTTTGAGCGGATA 240

Db

181 ATACCCACATTCGGGCCAGAGAGGCTTGGAGTTGGCTCTCAGCAGCTTTGAGCGGATA 240

Qy

241 AACAGGAAGACCTGTGCGGAGAGAGGACAGAGAGGACCTGTGTGAGGATACCCACCT 288

Db

241 AACAGGAAGACCTGTGCGGAGAGAGGACAGAGAGGACCTGTGTGAGGATACCCACCT 300

Qy

289 ----- 288

Db

301 GGTGGCCGCTCTCACTTGGGAACCACTCAACATGCGCTTCTTGGAACTCTCTTGTCACT 360

Qy

289 -----GATCCCAAGAACTTACAGGAGCTATGTCCGAGGAAATTCGGGCTCATG 339

Db

361 CCAAGAAAGATCCCAAGAACTTACAGGAGCTATGTCCGAGGAAATTCGGGCTCATG 420

Qy

340 GAAGACCGCAATGCGCGCTTAGGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTC 399

Db

421 GAAGACCGCAATGCGCGCTTAGGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTC 480

Qy

400 CTGCTGTGTAAGAGGACCTCAAAACCCCATGCGGTCCAGAGCAGCTTCTGGAACAGGC 459

Db

481 CTGCTGTGTAAGAGGACCTCAAAACCCCATGCGGTCCAGAGCAGCTTCTGGAACAGGC 540

Qy

460 CGGGGACACGGAGGACCTGGGACACCAAGCTAGCCCATCAGATAGAGACCTCTTT 519

Db

541 CGGGGACACGGAGGACCTGGGACACCAAGCTAGCCCATCAGATAGAGACCTCTTT 600

520 GAGCCAGAGAGAGCGCCCGAGCCACCGCGACCGTGGTTCATGCAAGGCGCGGAGGG 579
501 GAGCCAGAGAGAGCGCCCGAGCCACCGCGACCGTGGTTCATGCAAGGCGCGGAGGG 660
580 ATAGGCAAGTCCATGTGGGACACAAGAGTGTAGCTGGAGCTGGCGGAGCGGAGGCTCTTC 639
661 ATAGGCAAGTCCATGTGGGACACAAGAGTGTAGCTGGAGCTGGCGGAGCGGAGGCTCTTC 720
640 CAGGCGAGATTCATATATCTCTCTCATCACTGAGGAGAGATGAACAGAGTGCACAG 699
721 CAAGGCGAGATTCATATATCTCTCTCATCACTGAGGAGAGATGAACAGAGTGCACAG 780
700 GAATGCGAGATTCGCAAGACCTCATCTTCAGCTGTGGCTGAGCCGAGCGGCTCTCCAG 759
781 GAATGCGAGATTCGCAAGACCTCATCTTCAGCTGTGGCTGAGCCGAGCGGCTCTCCAG 840
760 GAGCTCATCGAGTTCGCGAGCGGCTCTCTCATCATGACGCTTCGATGAGCTCAAG 819
841 GAGCTCATCGAGTTCGCGAGCGGCTCTCTCATCATGACGCTTCGATGAGCTCAAG 900
820 CTTCTTTTCCACGATCTCAGGAGACCTGTGGCTCTGTGGAGAGAGAAACGCGCCACG 879
901 CTTCTTTTCCACGATCTCAGGAGACCTGTGGCTCTGTGGAGAGAGAAACGCGCCACG 960
880 GAGCTGCTTTTAAACAGCTTAATTCGGAAGAGCTGCTCCTGAGCTATCTTTGCTCATC 939
961 GAGCTGCTTTTAAACAGCTTAATTCGGAAGAGCTGCTCCTGAGCTATCTTTGCTCATC 1020
940 ACCACGCGCCACGCTTTGGAGAGCTCCACCGCTCTGTGGAGACCCCGGCGATGTG 999
1021 ACCACGCGCCACGCTTTGGAGAGCTCCACCGCTCTGTGGAGACCCCGGCGATGTG 1080
1000 GAGATCCTGGGCTTCTGAGGCGAAGAGAGAAATCTTCTACAGTATTTCCCAAT 1059
1081 GAGATCCTGGGCTTCTGAGGCGAAGAGAGAAATCTTCTACAGTATTTCCCAAT 1140
1060 GCAGAGAGGCGGCGGAGCTTCAATTAATCTGAGGAGACAGGAGCTCTCTCACCAGT 1119
1141 GCAGAGAGGCGGCGGAGCTTCAATTAATCTGAGGAGACAGGAGCTCTCTCACCAGT 1200
1120 TGCCTTCCTCCCTGCTGTGGTGTGTAGCTGCTTCGAGAGAGCTGGAGGGT 1179
1201 TGCCTTCCTCCCTGCTGTGGTGTGTAGCTGCTTCGAGAGAGCTGGAGGGT 1260
1180 GGGGGGCTGTTGAGAGAGAGCTCCAGGACCACTGCTGAGTGTATCTCTCAGAGAGCT 1239
1261 GGGGGGCTGTTGAGAGAGAGCTCCAGGACCACTGCTGAGTGTATCTCTCAGAGAGCT 1320
1240 AGTCTGATGCAACCCAGCGGGGCGCGCGCTCCAGCCCGCCACCCAGAGAGGG 1299
1321 AGTCTGATGCAACCCAGCGGGGCGCGCGCTCCAGCCCGCCACCCAGAGAGGG 1380
1300 TTGTGCTCTTGGCGGAGATGGGCTTGGAATCAGAAATCTTATTTGAGGAGGAGAC 1359
1381 TTGTGCTCTTGGCGGAGATGGGCTTGGAATCAGAAATCTTATTTGAGGAGGAGAC 1440
1360 CTCCGAGAGCGGCTGAGCGGGAGAGCTCTCTGCTCTCTCAACATGAATCTCTC 1419
1441 CTCCGAGAGCGGCTGAGCGGGAGAGCTCTCTGCTCTCTCAACATGAATCTCTC 1500
1420 CAGAAGGACATCAACTGTGAGAGGTACTACAGCTTCATCCATGAGTTTCAGGAATTC 1479
1501 CAGAAGGACATCAACTGTGAGAGGTACTACAGCTTCATCCATGAGTTTCAGGAATTC 1560
1480 TTGCGAGTATGATCTATATCTTCGAGCGGGGAGGCGGCGGCGGCGGCGGCGGAGAC 1539
1561 TTGCGAGTATGATCTATATCTTCGAGCGGGGAGGCGGCGGCGGCGGCGGCGGAGAC 1620
1540 GTGACAGGCTGTGACCGAGTACGGCTTTTCTGAAAGAGGCTTCCTGCGCATCTCACCAG 1599
1621 GTGACAGGCTGTGACCGAGTACGGCTTTTCTGAAAGAGGCTTCCTGCGCATCTCACCAG 1680
1600 CGCTTCTGTTGGACTCTCTGAAGAGGAGACCCAGGAGCCACTGGAGAGAGTCTCTGC 1659

1681 CGCTTCTGTTGGACTCTCTGAACGAGGAGACCCAGGAGGACCTGCGAGAGAGTCTCTGC 1740
1660 TGGAGGCTCTGCGCGCATCAAGATGAGACTTGTTCAGTGGATCCAAAGCAAGCTCAG 1719
1741 TGGAGGCTCTGCGCGCATCAAGATGAGACTTGTTCAGTGGATCCAAAGCAAGCTCAG 1800
1720 AGCGAGGCTCCACCTGCGAGGAGGCTCTCTGAGTTCCTTCAGCTGCTTGTACGAGATC 1779
1801 AGCGAGGCTCCACCTGCGAGGAGGCTCTCTGAGTTCCTTCAGCTGCTTGTACGAGATC 1860
1780 CAGGAGGAGGAGTTTATCCAGCAGGCGCTGAGCCATCTTCAGGTGATCGTGGTCAGCAAC 1839
1861 CAGGAGGAGGAGTTTATCCAGCAGGCGCTGAGCCATCTTCAGGTGATCGTGGTCAGCAAC 1920
1840 ATTGCTCCAGATGAGGAGCATGCTCTCTGCTTCTGTCTGAGCGCTGCGAGGAGCGCC 1899
1921 ATTGCTCCAGATGAGGAGCATGCTCTCTGCTTCTGTCTGAGCGCTGCGAGGAGCGCC 1980
1900 CAGGTGCTGCACTTGTATGGCGCCACCTACAGCGCGAGCGGGGAGAACCGCGCGAGGTGC 1959
1981 CAGGTGCTGCACTTGTATGGCGCCACCTACAGCGCGAGCGGGGAGAACCGCGCGAGGTGC 2040
1960 TCCGAGGAGCGCACACGCTGTTGGTGCAGCTCAGACGAGAGAGCCGTTCTCTCGAC 2019
2041 TCCGAGGAGCGCACACGCTGTTGGTGCAGCT---ACCAGAGAGGACCGTTCCTCTCGAC 2097
2020 GCTACAGTGAACATCTGCGCAGCGGCGCTGTGCGACCAATCCAAACCTGATAGAGTGTCT 2079
2098 GCTACAGTGAACATCTGCGCAGCGGCGCTGTGCGACCAATCCAAACCTGATAGAGTGTCT 2157
2080 CTGTACCGAAATTCGCTGCGGAGCGCGGGGGTGAAGTCTCTGTCAGGAGTCTCAGACAC 2139
2158 CTGTACCGAAATTCGCTGCGGAGCGCGGGGGTGAAGTCTCTGTCAGGAGTCTCAGACAC 2217
2140 CCCAATGCAAACTTCAGAACTGAGGCTGGAAGAGTCCGCGCATCTCCAGCTCAGCCTGC 2199
2218 CCCAATGCAAACTTCAGAACTGAGGCTGGAAGAGTCCGCGCATCTCCAGCTCAGCCTGC 2277
2200 GAGGAGCTCTCTGCGAGCTCTCATAGCCAAATGAAGATTTGACAGGATGGATCTCAGTGGC 2259
2278 GAGGAGCTCTCTGCGAGCTCTCATAGCCAAATGAAGATTTGACAGGATGGATCTCAGTGGC 2337
2260 AAGCGGCTTGGATTCAGGAGCATGCTGCTTTCCGAGGGGCTCGCGCATCCCGAGTGC 2319
2338 AAGCGGCTTGGATTCAGGAGCATGCTGCTTTCCGAGGGGCTCGCGCATCCCGAGTGC 2397
2320 AGGCTGAGATGATTCAGTGGAGAGTGTGAGTGGAGTCCGGGCTTGTGAGGAGATG 2379
2398 AGGCTGAGATGATTCAGTGGAGAGTGTGAGTGGAGTCCGGGCTTGTGAGGAGATG 2457
2380 GCTTCTGTGCTGGGACCAACCCACATCTGGTGGATGGAGCTGACAGGAAATGCATCTG 2439
2458 GCTTCTGTGCTGGGACCAACCCACATCTGGTGGATGGAGCTGACAGGAAATGCATCTG 2517
2440 GAGGATTTGGGCTGAGGTTTACTATCCAGGGGACTGAGGACCCAGTCTGCGAGTACCG 2499
2518 GAGGATTTGGGCTGAGGTTTACTATGCGAGGAGTCTGAGGACCCAGTCTGCGAGTACCG 2577
2500 ACTTTGTGGCTGAAGATCTGCCGCTCATCTGCTGCTGCTGAGCTGAGTGGGCTCAACT 2559
2578 ACTTTGTGGCTGAAGATCTGCCGCTCATCTGCTGCTGCTGAGCTGAGTGGGCTCAACT 2637
2560 CTGAGTGTGAACAGAGGCTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2619
2638 CTGAGTGTGAACAGAGGCTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2697
2620 GGGGTGCTGCTGCTGCTGAGGCGCTCAGGCAATCCCAAGTCCAGAGCTCCAGAGCTCCGCG 2679
2698 GGGGTGCTGCTGCTGCTGAGGCGCTCAGGCAATCCCAAGTCCAGAGCTCCAGAGCTCCGCG 2757
2680 TTGGGCACTGCGCGGCTGGGCTCTGCGCGCTGAGGCTCTTCTGCTGCTGCTCCAGGCGC 2739

Db 2758 TTGGGATCTGCGGCTGCTGCGCTGTGAGGCTCTTCTGTGGTCTCAGGCC 2817
Qy 2740 AACCAAACTCCGGAGCTGGACTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGG 2799
Db 2818 AACCAAACTCCGGAGCTGGACTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGG 2877
Qy 2800 TTGCTGGTGAAGGCTGCAATCTCCGCTGCGAGCTCCAGAACTGTGGCTGGATAGC 2859
Db 2878 TTGCTGGTGAAGGCTGCAATCTCCGCTGCGAGCTCCAGAACTGTGGCTGGATAGC 2937
Qy 2860 TGTGGCTTCACAGCCAAAGCTTGTGAGATCTTTACTTTCACCTGGGGATCAACGAGCC 2919
Db 2938 TGTGGCTTCACAGCCAAAGCTTGTGAGATCTTTACTTTCACCTGGGGATCAACGAGCC 2997
Qy 2920 TTGACCGACTTTACTTGACCAACAAACGCTTAGGGACACAGGTGTCCGACTGCTTTGC 2979
Db 2998 TTGACCGACTTTACTTGACCAACAAACGCTTAGGGACACAGGTGTCCGACTGCTTTGC 3057
Qy 2980 AAGGGCTGAGCATCTCGGCTGCAACTCCGAGTCTCTGTTTATTTGGATGAGACTG 3039
Db 3058 AAGGGCTGAGCATCTCGGCTGCAACTCCGAGTCTCTGTTTATTTGGATGAGACTG 3117
Qy 3040 AATAAATGACCCAGTAGGTGCGAGCGCTTCGAGTAACAAACCTTATTTGACATT 3099
Db 3118 AATAAATGACCCAGTAGGTGCGAGCGCTTCGAGTAACAAACCTTATTTGACATT 3177
Qy 3100 GGCTGCTGA 3108
Db 3178 GGCTGCTGA 3186

RESULT 6

ADF94754
ID ADF94754 standard; cDNA; 3186 BP.

XX ADF94754;

AC ADF94754;

DT 26-FEB-2004 (first entry)

XX Human PYRIN-8 coding sequence #3.

KW human; PYRIN; inflammatory disorder; inappropriate apoptosis;
KW inflammatory bowel disease; rheumatoid arthritis; diabetes;
KW multiple sclerosis; Grave's disease; contact dermatitis; psoriasis;
KW graft rejection; asthma; allergy; chronic obstructive pulmonary disease;
KW glomerulonephritis; infection; Alzheimer's disease; Parkinson's disease;
KW anaemia; ischaemia; screening; chromosomal mapping; tissue typing;
KW forensic biology; pharmacogenomics; predictive medicine; gene; ss.

XX Homo sapiens.

OS WO2003089588-A2.

XX 30-OCT-2003.

XX 14-APR-2003; 2003WO-US011572.

XX 17-APR-2002; 2002US-00124498.

XX (MILL-) MILLENNIUM PHARM INC.

FA (AMHP) MYETH.

XX Bertin J, Wang W, Blatcher M;

XX WPI; 2003-845527/78.

DR P-PSDB; ADF94755.

XX New nucleic acid molecules and polypeptides (e.g. PYRIN-2 or PYRIN-3)
XX useful for diagnosing, preventing or treating inflammation or disorders
XX associated with inappropriate apoptosis, in chromosomal mapping or in
XX pharmacogenomics.

PS Claim 4; SEQ ID NO 17; 199pp; English.

XX

The invention comprises the amino acid and coding sequences of human PYRIN proteins. The DNA and protein sequences of the invention are useful in diagnosing, preventing and treating inflammatory disorders or disorders associated with inappropriate apoptosis, such as: inflammatory bowel disease, rheumatoid arthritis, diabetes, multiple sclerosis, Grave's disease, contact dermatitis, psoriasis, graft rejection, asthma, allergy, chronic obstructive pulmonary disease, glomerulonephritis, infections, Alzheimer's disease, Parkinson's disease, anaemia and ischaemia. The DNA and protein sequences of the invention may also be used in screening assays, chromosomal mapping, tissue typing, forensic biology, pharmacogenomics, predictive medicine, and in monitoring of clinical trials. The present DNA sequence encodes a PYRIN protein of the invention.

XX Sequence 3186 BP; 701 A; 908 C; 945 G; 632 T; 0 U; 0 Other;

Query Match 96.5%; Score 2997.8; DB 10; Length 3186;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 3103; Conservative 0; Mismatches 2; Indels 84; Gaps 2;

Qy 1 ATGCTACGAAACCGCAGCAGCGGCTCTGTGCGCTGTCCACTACTTGGAGAACTC 60
Db 1 ATGCTACGAAACCGCAGCAGCGGCTCTGTGCGCTGTCCACTACTTGGAGAACTC 60
Qy 61 GAGGCTGTGGAACCTGAAGAAAGTTCAAGTTATACCTGGGGACCGCGACAGAGCTGGAGAA 120
Db 61 GAGGCTGTGGAACCTGAAGAAAGTTCAAGTTATACCTGGGGACCGCGACAGAGCTGGAGAA 120
Qy 121 GCGAAGATCCCTGGGGAGCATGGAAGAGCGGCTCCCTGGAAATGGCCAGCTGCTC 180
Db 121 GCGAAGATCCCTGGGGAGCATGGAAGAGCGGCTCCCTGGAAATGGCCAGCTGCTC 180
Qy 181 ATCACCACATTCGGGCCAGAGAGGCTGAGGTTGGCTCTCAGACACTTTGAGCGGATA 240
Db 181 ATCACCACATTCGGGCCAGAGAGGCTGAGGTTGGCTCTCAGACACTTTGAGCGGATA 240
Qy 241 AACAGGAAGACCTGTGGGAGAGAGGACAGAGAGAGGACCTGTGTGAGG----- 288
Db 241 AACAGGAAGACCTGTGGGAGAGAGGACAGAGAGAGGACCTGTGTGAGGATACCCACCT 300
Qy 289 ----- 288
Db 301 GGTGGCCGCTCTCACTTGGGAAACAGTCAACATGCCCTCTTGGAAAGTCTCTCTTGTCACT 360
Qy 289 -----CATCCCGAGAAACCTACAGGACTATGTCCGCGAGGAATTCGGGCTCATG 339
Db 361 CCAAGAAAGATCCCGAGGAAACCTACAGGACTATGTCCGCGAGGAATTCGGGCTCATG 420
Qy 340 GAAGACCGCAATGCGGCTTAGGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTC 399
Db 421 GAAGACCGCAATGCGGCTTAGGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTC 480
Qy 400 CTGCTGTGAAGAGAGACTCAAAACCCCATCGAGTCCAGCAGCAGCTTCTGGACACAGGC 459
Db 481 CTGCTGTGAAGAGAGACTCAAAACCCCATCGAGTCCAGCAGCAGCTTCTGGACACAGGC 540
Qy 460 CGGGGACACCGGAGGACCGTGGGACACCCAGCTAGCCCATCAAGTAGAGACCTCTTT 519
Db 541 CGGGGACACCGGAGGACCGTGGGACACCCAGCTAGCCCATCAAGTAGAGACCTCTTT 600
Qy 520 GAGCCAGACGAGGAGGCGCCCGGAGCCACCGCGCACCGTGTGTATGCAAGCGCGGACGGG 579
Db 601 GAGCCAGACGAGGAGGCGCCCGGAGCCACCGCGCACCGTGTGTATGCAAGCGCGGACGGG 660
Qy 580 ATAGGCAAGTCCATGTGGCACACAAGGTGATGCTGGACTGGGCGGACGGAAGCTTTC 639
Db 661 ATAGGCAAGTCCATGTGGCACACAAGGTGATGCTGGACTGGGCGGACGGAAGCTTTC 720
Qy 640 CAAGGCAAGTTCATTTATCTTCTACATCAACTGCGAGGAGATGACACGAGTGCACG 699
Db 721 CAAGGCAAGTTCATTTATCTTCTACATCAACTGCGAGGAGATGAAACGAGAGTGCCACG 780

Db 2938 TGTGGCTCAGCAGCAAGGCTTGTGAGAACTTTTACTTCACTCCCTGGGGATCAACAGACC 2997
Qy 2920 TTGACCGACCTTTACTGACCAACAACGCCCTAGGGGACACAGGTGTGCCGACTGCTTTGC 2979
Db 2998 TTGACCGACCTTTACTGACCAACAACGCCCTAGGGGACACAGGTGTGCCGACTGCTTTGC 3057
Qy 2980 AAGGGCTGAGCATCTCTGGCTGCAAACTCCGAGTCTCTGGTTATTTGGGATGAGACTG 3039
Db 3058 AAGGGCTGAGCATCTCTGGCTGCAAACTCCGAGTCTCTGGTTATTTGGGATGAGACTG 3117
Qy 3040 AATAAATGACCCACAGTAGCTTGCAGCGCTTCGAGTAAACAACCTTATTTCGACATT 3099
Db 3118 AATAAATGACCCACAGTAGCTTGCAGCGCTTCGAGTAAACAACCTTATTTCGACATT 3177
Qy 3100 GGCTGCTGA 3108
Db 3178 GGCTGCTGA 3186

RESULT 7
ADJ19382
ID ADJ19382 standard; cDNA; 3218 BP.
XX
AC ADJ19382;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human PAN6 PAAD domain-containing protein full-length cDNA 2.
XX
KW PAAD; cytostatic; cancer; gene therapy; human; PAAD domain; ss; gene;
KW PAN6.
XX
OS Homo sapiens.
XX
PN US2004002593-A1.
XX
PD 01-JAN-2004.
XX
PF 04-APR-2003; 2003US-00407866.
XX
PR 04-APR-2002; 2002US-0370538P.
XX
PA (REED/) REED J C.
FA (GODZ/) GODZIK A.
XX
PI Reed JC, Godzik A;
XX
DR WPI; 2004-061677/06.
XX
P-PSDB; ADJ19383.
XX
PT New nucleic acid molecule encoding a PAAD-containing polypeptide, useful
PT for preparing a composition for diagnosing or treating e.g., cancer.
XX
PS Disclosure; SEQ ID NO 67; 193pp; English.
XX
CC The invention relates to a novel isolated nucleic acid molecule encoding
CC a PAAD-containing polypeptide. The polypeptide of the invention
CC demonstrates cytostatic activity and may be useful for preparing a
CC composition for diagnosing or treating diseases associated with the PAAD
CC domain-containing polypeptide, such as cancer, via gene therapy. The
CC current sequence is that of the human PAAD domain-containing protein full
CC -length cDNA of the invention.
SQ Sequence 3218 BP; 691 A; 935 C; 958 G; 634 T; 0 U; 0 Other;

Query Match 94.2%; Score 2929; DB 12; Length 3218;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 3020; Conservative 0; Mismatches 0; Indels 81; Gaps 1;
Qy 1 ATGCTACGAACCGCAGGACGAGCGGCTCTGTGCGCTGTCCACCTACTTGGAGAATCTC 60
Db 118 ATGCTACGAACCGCAGGACGAGCGGCTCTGTGCGCTGTCCACCTACTTGGAGAATCTC 177

Qy 61 GAGGCTGTGAACTGAAGAAAGTTCAAGTTATACCTGGGACCGCGACAGAGCTGGAGAA 120
Db 178 GAGGCTGTGAACTGAAGAAAGTTCAAGTTATACCTGGGACCGCGACAGAGCTGGAGAA 237
Qy 121 GGCAGATCCCTGGGGAAGCATGGAGGCGGCTCCCTGGAAATGGCCAGCTGCTC 180
Db 238 GGCAGATCCCTGGGGAAGCATGGAGGCGGCTCCCTGGAAATGGCCAGCTGCTC 297
Qy 181 ATCACCACCTTCGGGCGCAGAGGAGGCTGGAGGTTGGCTCTCAGCACCTTTGAGCGGATA 240
Db 298 ATCACCACCTTCGGGCGCAGAGGAGGCTGGAGGTTGGCTCTCAGCACCTTTGAGCGGATA 357
Qy 241 AACAGGAAGACCTGTGGGAGAGGAGACAGAGAGAGGACCTGTGTGAGG- 288
Db 358 AACAGGAAGACCTGTGGGAGAGGAGACAGAGAGAGGACCTGTGTGAGGAGTACCCCACT 417
Qy 289 ----- 288
Db 418 GGTGGCCGCTCTCATTGGGAAACCATGTCATGCTTCTGGAAAGTCTCTCTTGTCACT 477
Qy 289 -----GATCCCCAGGAAACCTACAGGAGACTATGTCCGCGAGGAAATTCGGCTCATG 339
Db 478 CCAAGAAAAGATCCCCAGGAAACCTACAGGAGCTATGTCCGCGAGGAAATTCGGCTCATG 537
Qy 340 GAAGACCGCAATGCGGCCCTAGGGGAATGTCAAACCTCAGGCCACCGGTACACCGGCTC 399
Db 538 GAAGACCGCAATGCGGCCCTAGGGGAATGTCAAACCTCAGGCCACCGGTACACCGGCTC 597
Qy 400 CTGCTGTGAGGAGGACCTCAAAACCCCATGAGGTCCAGGAGGAGCTTCTGGACACAGGC 459
Db 598 CTGCTGTGAGGAGGACCTCAAAACCCCATGAGGTCCAGGAGGAGCTTCTGGACACAGGC 657
Qy 460 CGGGGACACGCGAGGACCGTGGGACACACGAGCTAGCCCCATCAAGATAGAGACCTCTTT 519
Db 658 CGGGGACACGCGAGGACCGTGGGACACACGAGCTAGCCCCATCAAGATAGAGACCTCTTT 717
Qy 520 GAGCCAGACGAGGAGCGCCCGAGCCACCGGCAACCGTGTCTATGCAAGGCGCGGACGG 579
Db 718 GAGCCAGACGAGGAGCGCCCGAGCCACCGGCAACCGTGTCTATGCAAGGCGCGGACGG 777
Qy 580 ATAGGCAAGTCCATGCTGGCACAACAGGTGATGCTGGACTGGGCGGACGGGAAGCTCTTC 639
Db 778 ATAGGCAAGTCCATGCTGGCACAACAGGTGATGCTGGACTGGGCGGACGGGAAGCTCTTC 837
Qy 640 CAAGGCAGATTTGATTATCTCTTACATCAACTGACGAGGAGATGAACAGAGTGCCACG 699
Db 838 CAAGGCAGATTTGATTATCTCTTACATCAACTGACGAGGAGATGAACAGAGTGCCACG 897
Qy 700 GAATGACAGCATGCAAGACCTCATCTTCAGCTGTGGCTGAGCCGAGCGGCTCTCCAG 759
Db 898 GAATGACAGCATGCAAGACCTCATCTTCAGCTGTGGCTGAGCCGAGCGGCTCTCCAG 957
Qy 760 GAGCTCATCGAGTTCCCGAGCGCTCTTTTATCATCATCGAGCTTCGATGAGCTCAAG 819
Db 958 GAGCTCATCGAGTTCCCGAGCGCTCTTTTATCATCATCGAGCTTCGATGAGCTCAAG 1017
Qy 820 CTTCTTTTCCAGCATCTCAGGAGACCTGCTGCTCTGCTGGAGGAGAAACGCCCAACG 879
Db 1018 CTTCTTTTCCAGCATCTCAGGAGACCTGCTGCTCTGCTGGAGGAGAAACGCCCAACG 1077
Qy 880 GAGCTGTTCTTAAACAGCTTAAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATC 939
Db 1078 GAGCTGTTCTTAAACAGCTTAAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATC 1137
Qy 940 ACCACAGGCGCCAGGCTTTGGAGAGCTCCACGCTCTGCTGGAGCACCCAGGCAATGTG 999
Db 1138 ACCACAGGCGCCAGGCTTTGGAGAGCTCCACGCTCTGCTGGAGCACCCAGGCAATGTG 1197
Qy 1000 GAGATCTGGGCTTCTCTGAGGAGAGAAAGGAATACCTTCTACAGTATTTCCAAAT 1059
Db 1198 GAGATCTGGGCTTCTCTGAGGAGAGAAAGGAATACCTTCTACAGTATTTCCAAAT 1257
Qy 1060 GCAGAGCAGCGGGCCAAAGTCTTTCAATTACGTGAGGAGCAACGAGGCTCTCTTCCACCATG 1119

| | | | |
|----------|--|--|------|
| Db | 2801 | ACTTTGTGGCTGAAGATCTGCGCCTCACTGCTGCTGCTGACGAGCTGGCTCAACT | 2869 |
| Qy | 2560 | CTCAGTGTGAACAGAGCCTGAGAGAGCTGGACCTGAGCCTGAATGAGCTGGGGGACCTC | 2619 |
| Db | 2861 | CTCAGTGTGAACAGAGCCTGAGAGAGCTGGACCTGAGCCTGAATGAGCTGGGGGACCTC | 2920 |
| Qy | 2620 | GGGTGTGCTGCTGTGTGTGAGGGCCTCAGGCATCCCACTGCAAGCTCCAGACCCTCGG | 2679 |
| Db | 2921 | GGGTGTGCTGCTGTGTGTGAGGGCCTCAGGCATCCCACTGCAAGCTCCAGACCCTCGG | 2980 |
| Qy | 2680 | TTGGGCATCTGCGCGCTGGGCTCTGCGCCTGTGAGGGTCTTTCTGTGTGCTCCAGGCC | 2739 |
| Db | 2981 | TTGGGCATCTGCGCGCTGGGCTCTGCGCCTGTGAGGGTCTTTCTGTGTGCTCCAGGCC | 3040 |
| Qy | 2740 | AACCACAACTCTCCGGAGCTGGACTTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGG | 2799 |
| Db | 3041 | AACCACAACTCTCCGGAGCTGGACTTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGG | 3100 |
| Qy | 2800 | TTGCTGCTGAGGGGCTGCAACATCCCGCCTGCAGACTCCAGAACTGTGG | 2850 |
| Db | 3101 | TTGCTGCTGAGGGGCTGCAACATCCCGCCTGCAGACTCCAGAACTGTGG | 3151 |
| RESULT 9 | | | |
| ADM02033 | | | |
| ID | ADM02033 standard; cDNA; 3466 BP. | | |
| XX | | | |
| AC | ADM02033; | | |
| XX | | | |
| DT | 20-MAY-2004 (first entry) | | |
| DE | Human cDNA of the invention SEQ ID NO:718. | | |
| DE | ss; gene; human; gene therapy; diagnostic marker; pharmaceutical. | | |
| KW | Homo sapiens. | | |
| XX | | | |
| OS | EP1347046-A1. | | |
| XX | | | |
| PN | 24-SEP-2003. | | |
| XX | | | |
| PD | 12-APR-2002; 2002EP-00008400. | | |
| XX | | | |
| PF | 22-MAR-2002; 2002JP-00137785. | | |
| XX | | | |
| PR | (REAS-) RES ASSOC BIOTECHNOLOGY. | | |
| XX | | | |
| PA | Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S; | | |
| XX | Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I; | | |
| PI | Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuho Y; | | |
| PI | WPI; 2003-723558/69. | | |
| XX | P-PSDB; ADM04476. | | |
| DR | | | |
| DR | | | |
| XX | New polynucleotides and polypeptides are useful in gene therapy, for | | |
| PT | developing a diagnostic marker or medicines for regulating their | | |
| PT | expression and activity, or as a target of gene therapy. | | |
| XX | | | |
| XX | Claim 1; SEQ ID NO 718; 305pp; English. | | |
| PS | | | |
| XX | The invention relates to a novel human polynucleotide and the encoded | | |
| CC | polypeptide. A polynucleotide of the invention may have a use in gene | | |
| CC | therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful | | |
| CC | as a primer for synthesizing the polynucleotide or as a probe for | | |
| CC | detecting the polynucleotide. The polynucleotides ADM03116-ADM03758 are | | |
| CC | useful in gene therapy, for developing a diagnostic marker or medicines | | |
| CC | for regulating their expression and activity, or as a target of gene | | |
| CC | therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides | | |
| CC | are useful as pharmaceutical agents. The present sequence represents a | | |
| CC | cDNA sequence of the invention. | | |
| XX | | | |
| XX | Sequence 3466 BP; 780 A; 985 C; 1005 G; 696 T; 0 U; 0 Other; | | |
| SQ | | | |

| | | | | | |
|-----------------------|------|--|---------------|------------|--------------|
| Query Match | | 85.1%; | Score 2644.2; | DB 11; | Length 3466; |
| Best Local Similarity | | 91.9%; | Pred. No. 0; | | |
| Matches 2931; | | Conservative | 0; | Mismatches | 3; |
| | | | | Indels | 255; |
| | | | | Gaps | 3; |
| Qy | 1 | ATGCTACGAACCGCAGGCGAGGACGGCCCTCTGTGCCCTGTCCACTTCTTGGAAGACTC | 60 | | |
| Db | 145 | ATGCTACGAACCGCAGGCGAGGACGGCCCTCTGTGCCCTGTCCACTTCTTGGAAGACTC | 204 | | |
| Qy | 61 | GAGGCTGTGGAACTGAAGAAGTTCAAGTTATACCTTGGGGACCGGACACAGAGCTGGAGAA | 120 | | |
| Db | 205 | GAGGCTGTGGAACTGAAGAAGTTCAAGTTATACCTTGGGGACCGGACACAGAGCTGGAGAA | 264 | | |
| Qy | 121 | GGCAAGATCCCTCGGGGAAGCATGGAGAAAGCCCGGTCCCTTGGAAATGGCCCGAGCTGCTC | 180 | | |
| Db | 265 | GGCAAGATCCCTCGGGGAAGCATGGAGAAAGCCCGGTCCCTTGGAAATGGCCCGAGCTGCTC | 324 | | |
| Qy | 181 | ATCACCCACTTTCGGCCACAGAGAGCCCTGGAGGTTGGCTCTCAGCACCTTTGAGGGATA | 240 | | |
| Db | 325 | ATCACCCACTTTCGGCCACAGAGAGCCCTGGAGGTTGGCTCTCAGCACCTTTGAGGGATA | 384 | | |
| Qy | 241 | AACAGGAAGACCTGTGTGGACAGAGGACAGAGAGAGACCTGGTGAGG----- | 288 | | |
| Db | 385 | AACAGGAAGACCTGTGTGGACAGAGGACAGAGAGAGACCTGGTGAGGATACCCCACCT | 444 | | |
| Qy | 289 | ----- | 288 | | |
| Db | 445 | GGTGGCCCGTCTCACTTGGGAACCCAGTCAACATGSCCTTCTTGGAAAGTCTCTTTGTCACT | 504 | | |
| Qy | 289 | -----GATCCCCAGGAACCTACAGGAGACTATGTCCGACGAGAAATTCGGGCTCATG | 339 | | |
| Db | 505 | CCAAGAAAAGATCCCCAGGAACCTACAGGAGACTATGTCCGACGAGAAATTCGGGCTCATG | 564 | | |
| Qy | 340 | GAAGACCGCAATGGCGCTTAGGGGAATGTGCAACCTTCAGCACCCGCTACACCCGGCTC | 399 | | |
| Db | 565 | GAAGACCGCAATGGCGCTTAGGGGAATGTGCAACCTTCAGCACCCGCTACACCCGGCTC | 624 | | |
| Qy | 400 | CTGCTGTGTGAAGGACACTCAAAACCCCATGCAAGTCCAGCAGCAGCTTCTGGACACAGGC | 459 | | |
| Db | 625 | CTGCTGTGTGAAGGACACTCAAAACCCCATGCAAGTCCAGCAGCAGCTTCTGGACACAGGC | 684 | | |
| Qy | 460 | CGGGACACGGGAGACCGTGGGACACAGGCTAGCCCCCATCAAGATAGAGACCTCTTT | 519 | | |
| Db | 685 | CGGGGACACGGGAGACCGTGGGACACAGGCTAGCCCCCATCAAGATAGAGACCTCTTT | 744 | | |
| Qy | 520 | GAGCCAGACGAGGACGCCCGGACACCGGACCGGTGTCATGCAAGGCGCGGACGGG | 579 | | |
| Db | 745 | GAGCCAGACGAGGACGCCCGGACACCGGACCGGTGTCATGCAAGGCGCGGACGGG | 804 | | |
| Qy | 580 | ATAGGCAAGTCCATGCTGGACACAAAGGTGATGTGGACTGGCGGACGGGAAGCTCTTC | 639 | | |
| Db | 805 | ATAGGCAAGTCCATGCTGGACACAAAGGTGATGTGGACTGGCGGACGGGAAGCTCTTC | 864 | | |
| Qy | 640 | CAAGCAGATTTGATTATCTTCTATCACTGACGGAGATGAACAGAGTGCCACG | 699 | | |
| Db | 865 | CAAGCAGATTTGATTATCTTCTATCACTGACGGAGATGAACAGAGTGCCACG | 924 | | |
| Qy | 700 | GAATGCACATGCAAGACCTCATCTTCACTGCTGGCCCTGAGCCAGCGCCCTCTCCAG | 759 | | |
| Db | 925 | GAATGCACATGCAAGACCTCATCTTCACTGCTGGCCCTGAGCCAGCGCCCTCTCCAG | 984 | | |
| Qy | 760 | GAGCTCATCCGAGTTCCGAGCGCCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAG | 819 | | |
| Db | 985 | GAGCTCATCCGAGTTCCGAGCGCCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAG | 1044 | | |
| Qy | 820 | CCTTCTTTCCACGATTCCTCAGGACCCCTGGTGGCTCTGCTGGGAGAGAAACGGCCACG | 879 | | |
| Db | 1045 | CCTTCTTTCCACGATTCCTCAGGACCCCTGGTGGCTCTGCTGGGAGAGAAACGGCCACG | 1104 | | |
| Qy | 880 | GAGCTGCTTTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATC | 939 | | |
| Db | 1105 | GAGCTGCTTTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATC | 1164 | | |

| | | | | | |
|----|------|---|------|--|--|
| Qy | 940 | ACCAACGCGGCCACGGCTTTGGAGAAAGCTCCACCGTCTGTCTGGAGCACCCCGAGCATGTG | 999 | | |
| Db | 1165 | ACCAACGCGGCCACGGCTTTGGAGAAAGCTCCACCGTCTGTCTGGAGCACCCCGAGCATGTG | 1224 | | |
| Qy | 1000 | GAGATCTGGGCTTCTCTGAGGCGAAGGAAGGAATACTTCTCAAGTATTTCCACAAT | 1059 | | |
| Db | 1225 | GAGATCTGGGCTTCTCTGAGGCGAAGGAAGGAATACTTCTCAAGTATTTCCACAAT | 1284 | | |
| Qy | 1060 | GCAGAGCAGGCGGCGCCAAAGTCTTCAATTAAGTGAAGGACCAACGAGCCCTCTTCCACCATG | 1119 | | |
| Db | 1285 | GCAGAGCAGGCGGCGCCAAAGTCTTCAATTAAGTGAAGGACCAACGAGCCCTCTTCCACCATG | 1344 | | |
| Qy | 1120 | TGCTTCTGCTCCCTCTGGTGTGCTGGTGTGTAACCTGCTCCAGCAGCAGCTGAGGGT | 1179 | | |
| Db | 1345 | TGCTTCTGCTCCCTCTGGTGTGCTGGTGTGTAACCTGCTCCAGCAGCAGCTGAGGGT | 1404 | | |
| Qy | 1180 | GGGGGGCTGTTGAGACAGAGCTCCAGGACCAACCACTGACGTGATACATGCTCTACTGCTG | 1239 | | |
| Db | 1405 | GGGGGGCTGTTGAGACAGAGCTCCAGGACCAACCACTGACGTGATACATGCTCTACTGCTG | 1464 | | |
| Qy | 1240 | AGTCTGATGCAACCCAAAGCCGGGGCCCCCGGCTCCAGCCCCCAACCCAGAGAGGG | 1299 | | |
| Db | 1465 | AGTCTGATGCAACCCAAAGCCGGGGCCCCCGGCTCCAGCCCCCAACCCAGAGAGGG | 1524 | | |
| Qy | 1300 | TTGTGCTCTTTGGCGGAGATGGGCTCTGGAAATCAGAAATCTTATTTGAGGAGCAGGAC | 1359 | | |
| Db | 1525 | TTGTGCTCTTTGGCGGAGATGGGCTCTGGAAATCAGAAATCTTATTTGAGGAGCAGGAC | 1584 | | |
| Qy | 1360 | CTCCGGAAGCACCGCCCTAGACCGGGAAGAGCTCTGCTCTCCATCAATCAATCAATCTTC | 1419 | | |
| Db | 1585 | CTCCGGAAGCACCGCCCTAGACCGGGAAGAGCTCTGCTCTCCATCAATCAATCAATCTTC | 1644 | | |
| Qy | 1420 | CAGAAAGACATCAACTGTGAGAGGTACTACAGCTTTCATCCACTTTGAGTTTCCAGGAATTC | 1479 | | |
| Db | 1645 | CAGAAAGACATCAACTGTGAGAGGTACTACAGCTTTCATCCACTTTGAGTTTCCAGGAATTC | 1704 | | |
| Qy | 1480 | TTTGACGCTATGTAATACTTCCGGAAGGAGGAGGCGGGGAGGCGGCGGAGCAGGAC | 1539 | | |
| Db | 1705 | TTTGACGCTATGTAATACTTCCGGAAGGAGGAGGCGGGGAGGCGGCGGAGCAGGAC | 1764 | | |
| Qy | 1540 | GTGACACGAGCTGTGACCGGAGTACCGCTTTCTCAAGAGGAGCTTCTGGGACCTCAGCAGC | 1599 | | |
| Db | 1765 | GTGACACGAGCTGTGACCGGAGTACCGCTTTCTCAAGAGGAGCTTCTGGGACCTCAGCAGC | 1824 | | |
| Qy | 1600 | CGCTTCTCTGTTTGGACTCTCTGAACGAGGAGACCAAGGAGCCTTGGAGAAAGCTCTGCG | 1659 | | |
| Db | 1825 | CGCTTCTCTGTTTGGACTCTCTGAACGAGGAGACCAAGGAGCCTTGGAGAAAGCTCTGCG | 1884 | | |
| Qy | 1660 | TGGAAGGTCTCGCCGACATCAAGATGGAACCTGTTGAGTGGAATCCAAAGCAAAAGCTCAG | 1719 | | |
| Db | 1885 | TGGAAGGTCTCGCCGACATCAAGATGGAACCTGTTGAGTGGAATCCAAAGCAAAAGCTCAG | 1944 | | |
| Qy | 1720 | AGCGACGGCTCCACCTTGCAGCAGGGCTCCTTGGAGTTCTTCACTGCTGTCTGAGATC | 1779 | | |
| Db | 1945 | AGCGACGGCTCCACCTTGCAGCAGGGCTCCTTGGAGTTCTTCACTGCTGTCTGAGATC | 2004 | | |
| Qy | 1780 | CAGGAGGAGGAGTTTATCCAGCAGCGCCCTGAGCCTTCCAGGTCATCGTGGTCAGCAAC | 1839 | | |
| Db | 2005 | CAGGAGGAGGAGTTTATCCAGCAGCGCCCTGAGCCTTCCAGGTCATCGTGGTCAGCAAC | 2064 | | |
| Qy | 1840 | ATTGCCCTCCAAAGATGAGAGCAGTGGTCTCTCTGTTCTGTCTGAAAGCGCTGAGAGCGCC | 1899 | | |
| Db | 2065 | ATTGCCCTCCAAAGATGAGAGCAGTGGTCTCTCTGTTCTGTCTGAAAGCGCTGAGAGCGCC | 2124 | | |
| Qy | 1900 | CAGGTGCTGCACTTGTATGCGCCACCTACAGCGCGGAGCGGGAGACCGCGCAGGTCG | 1959 | | |
| Db | 2125 | CAGGTGCTGCACTTGTATGCGCCACCTACAGCGCGGAGCGGGAGACCGCGCAGGTCG | 2184 | | |
| Qy | 1960 | TCCGACGAGGCGCACGCTGTTGGTGCAGCTCAGACACAGACAGAGAGACCGTCTGCTGAC | 2019 | | |
| Db | 2185 | TCCGACGAGGCGCACGCTGTTGGTGCAGCT---ACCAGAGAGAGACCGTCTGCTGAC | 2241 | | |
| Qy | 2020 | GCCTACAGTGAACATCTGGCAGCGGCCCTGTGTGCAACCAATCCAAACCTGATAGAGCTGTCT | 2079 | | |

| | | | |
|----|------|---|------|
| Db | 2242 | GCCTACAGTGAACATCTGSCAGCGCCCTGTGACCAATCCAAACCTGATAGAGCTGTCT | 2301 |
| Qy | 2080 | CTGTACCGAAATGCCCTTGGGACGCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACAC | 2139 |
| Db | 2302 | CTGTACCGAAATGCCCTGCGGACGCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACAC | 2361 |
| Qy | 2140 | CCCAACTGCAAACTTCAGAAACCTGAGGCTGGAAGAGGTGCCGCATCTCCAGCTCAGCCCTGC | 2199 |
| Db | 2362 | CCCAACTGCAAACTTCAGAAACCTGAGGCTGGAAGAGGTGCCGCATCTCCAGCTCAGCCCTGC | 2421 |
| Qy | 2200 | GAGGACCTCTCTGACGCTCTCATAGCCAAATAAGAAATTTGACAAGGATGGATCTCAGTGGC | 2259 |
| Db | 2422 | GAGGACCTCTCTGACGCTCTCATAGCCAAATAAGAAATTTGACAAGGATGGATCTCAGTGGC | 2481 |
| Qy | 2260 | AACGGCGTTGGATTCCAGGCATGATGCTGCTTTCGAGGGGCTTCGGGCATCCCCAAATGC | 2319 |
| Db | 2482 | AACGGCGTTGGATTCCAGGCATGATGCTGCTTTCGAGGGGCTTCGGGCATCCCCAAATGC | 2541 |
| Qy | 2320 | AGGCTGCAGATGATTCAAGTTGAGGAAGTGTCAAGTCTGGAGTCCGGGGCTTGTCCAGAGATG | 2379 |
| Db | 2542 | AGGCTGCAGATGATTCAAGTTGAGGAAGTGTCAAGTCTGGAGTCCGGGGCTTGTCCAGAGATG | 2601 |
| Qy | 2380 | GCTTCTGTGCTCGGCACCAACCCACATCTGTGTTGAGTTGGAACCTGACAGGAAATGCACTG | 2439 |
| Db | 2602 | GCTTCTGTGCTCGGCACCAACCCACATCTGTGTTGAGTTGGAACCTGACAGGAAATGCACTG | 2661 |
| Qy | 2440 | GAGGATTTGGGCTGAGGTTACTATGCAAGGAACTGAGGCAACCGAGTCTGCAGACTACGG | 2499 |
| Db | 2662 | GAGGATTTGGGCTGAGGTTACTATGCAAGGAACTGAGGCAACCGAGTCTGCAGACTACGG | 2721 |
| Qy | 2500 | ACTTTGTGGCTGAAGATCTGCGGCTCAGCTGCTGCTGTGACAGCTGGCCTCAACT | 2559 |
| Db | 2722 | ACTTTGTGGCTGAAGATCTGCGGCTCAGCTGCTGCTGTGACAGCTGGCCTCAACT | 2781 |
| Qy | 2560 | CTCAGTGTGAACACAGAGCCTGAGAGACTGGACCTGAGCCTGAAATGAGCTGGGGGACCTC | 2619 |
| Db | 2782 | CTCAGTGTGAACACAGAGCCTGAGAGACTGGACCTGAGCCTGAAATGAGCTGGGGGACCTC | 2841 |
| Qy | 2620 | GGGTGTGCTGCTGTGTGTGAGGGGCTCAGAGCATCCACGTGCAAGCTCCAGACCCCTGCCG | 2679 |
| Db | 2842 | GGGTGTGCTGCTGTGTGTGAGGGGCTCAGAGCATCCACGTGCAAGCTCCAGACCCCTGCC | 2899 |
| Qy | 2680 | TTGGGCATCTCGCGCTGGGCTCTGCGCCCTGTGAGGGTCTTTCTGTGTGTCTCCAGGCC | 2739 |
| Db | 2900 | ----- | 2899 |
| Qy | 2740 | AACCAACCTCGGGAGCTGGACTTGAGTTTCAAACCTGGGAGACTGGGGCCTGTGG | 2799 |
| Db | 2900 | ----- | 2899 |
| Qy | 2800 | TTGCTGGCTGAGGGCTGCAACATCCCGCTTCAGACTCCAGAAACTGTGGCTGGATAGC | 2859 |
| Db | 2900 | -----GGCTGGATAGC | 2910 |
| Qy | 2860 | TGTGGCTTCACAGCAAGCTTGTGAGAACTTTTACTTCACTCCCTGGGGATCAACAGACC | 2919 |
| Db | 2911 | TGTGGCTTCACAGCAAGCTTGTGAGAACTTTTACTTCACTCCCTGGGGATCAACAGACC | 2970 |
| Qy | 2920 | TTTGACCGACTTTTACTTCAGCAAAACAAACCGCTTAGGGGACACAGGTGTCCGATGCTTTGC | 2979 |
| Db | 2971 | TTTGACCGACTTTTACTTCAGCAAAACAAACCGCTTAGGGGACACAGGTGTCCGATGCTTTGC | 3030 |
| Qy | 2980 | AAGCGCTGAGCAATCTGGCTGCAAACTCGAGTCTCTGGTTATTTGGGATGGACCTG | 3039 |
| Db | 3031 | AAGCGCTGAGCAATCTGGCTGCAAACTCGAGTCTCTGGTTATTTGGGATGGACCTG | 3090 |
| Qy | 3040 | AATAAATCACCAAGTAGGTGGCAGCGCTTCAGGTAAACAAACCTTATTTTGACATT | 3099 |
| Db | 3091 | AATAAATCACCAAGTAGGTGGCAGCGCTTCAGGTAAACAAACCTTATTTTGACATT | 3150 |
| Qy | 3100 | GGCTGCTGA 3108 | |

| | | | |
|-----------|----------|---|------|
| Db | 3151 | GGCTGCTGA | 3159 |
| RESULT 10 | | | |
| ADP47727 | | | |
| ID | ADP47727 | standard; cDNA; 3395 BP. | |
| XX | AC | ADP47727; | |
| XX | AC | | |
| XX | DT | 12-AUG-2004 (first entry) | |
| XX | XX | | |
| XX | XX | Human Monarch-1 (CATERPILLER 19.3) isoform III cDNA. | |
| XX | XX | Monarch-1; CATERPILLER 11.2; caspase recruitment domain; | |
| KW | KW | CARD, transcription enhancer, R-binding, pyrin, lots of leucine repeat; | |
| KW | KW | CATERPILLER 11.3; CATERPILLER 16.1; CATERPILLER 16.2; CIAS1; | |
| KW | KW | cold-induced autoinflammatory syndrome 1; antiinflammatory; cytosolic; | |
| KW | KW | inflammatory disease; cancer; gene therapy; human; gene; ss; purine; | |
| KW | KW | CATERPILLER 19.3; chromosome 19q13; multiple sclerosis; isoform III. | |
| XX | XX | | |
| OS | XX | Homo sapiens. | |
| XX | XX | | |
| XX | XX | WO2004034093-A2. | |
| XX | XX | | |
| PD | PD | 22-APR-2004. | |
| XX | XX | | |
| PF | XX | 30-APR-2003; 2003WO-US013562. | |
| XX | XX | | |
| PR | XX | 30-APR-2002; 2002US-0376626P. | |
| XX | XX | | |
| PA | XX | (UYNC-) UNIV NORTH CAROLINA. | |
| XX | XX | | |
| PI | XX | Ting JY, Linhoff WM, Harton JA, Williams KL, Lich J, O'connor W; | |
| PI | PI | Moore CB, Davis B, Brickey J, Conti B, Zhang J, Zhu X; | |
| XX | XX | | |
| DR | XX | WPI; 2004-348215/32. | |
| DR | XX | P-PSDB; ADP47728. | |
| XX | XX | | |
| PT | XX | New nucleic acid encoding Monarch-1, CATERPILLER 11.2, CATERPILLER 11.3, | |
| PT | XX | CATERPILLER 16.1, CATERPILLER 16.2 or CIAS1 polypeptide, useful in | |
| PT | XX | preparing a composition for treating inflammatory disease or cancer. | |
| XX | XX | | |
| PS | XX | Claim 4; SEQ ID NO 5; 205pp; English. | |
| XX | XX | | |
| CC | XX | The invention relates to a novel isolated nucleic acid encoding a Monarch | |
| CC | XX | -1, CATERPILLER (CARD [caspase recruitment domain], transcription | |
| CC | XX | enhancer, R(purine)-binding, pyrin, lots of leucine repeats) 11.2, | |
| CC | XX | CATERPILLER 11.3, CATERPILLER 16.1, CATERPILLER 16.2 or CIAS1 (cold- | |
| CC | XX | induced autoinflammatory syndrome 1) polypeptide comprising the amino | |
| CC | XX | acids encoded by exon 6 and lacking the amino acids encoded by exon 4 or | |
| CC | XX | its fragment. The nucleic acid of the invention demonstrates | |
| CC | XX | antiinflammatory and cytosolic activities and may be useful in preparing | |
| CC | XX | a composition for treating an inflammatory disease or cancer, possibly | |
| CC | XX | via gene therapy. The current sequence is that of the human Monarch-1 | |
| CC | XX | (CATERPILLER 19.3) isoform III cDNA of the invention which is located on | |
| CC | XX | chromosome 19q13, in the multiple sclerosis susceptibility region. | |
| XX | XX | | |
| SQ | XX | Sequence 3395 BP; 786 A; 946 C; 980 G; 683 T; 0 U; 0 Other; | |
| | | Query Match 79.5%; Score 2470; DB 12; Length 3395; | |
| | | Best Local Similarity 94.5%; Pred. No. 0; | |
| | | Matches 2616; Conservative 0; Mismatches 65; Indels 87; Gaps 2 | |
| QY | 1 | ATGCTACGAACCGCAGCGAGGACGGCCTCTGTCGCTGTCCACTCTTGGAGAACTC 60 | |
| Db | 221 | ATGCTACGAACCGCAGCGAGGACGGCCTCTGTCGCTGTCCACTCTTGGAGAACTC 280 | |
| QY | 61 | GAGGCTGTGGAACCTGGAAGAGTTCAAGTTATACCTGGGGACCGCAGACGAGCTGGAGAA 120 | |
| Db | 281 | GAGGCTGTGGAACCTGGAAGAGTTCAAGTTATACCTGGGGACCGCAGACGAGCTGGAGAA 340 | |
| QY | 121 | GGCAAGATCCCTGGGGAAGCATGGAGAGGCGCGTCCCTGGAAATGCCCCAGCTGCTC 180 | |

Db 341 GGCAGAGATCCCTGGGAGCAATGGAGAAAGCCCGTCCCTCGGAAATGGCCAGCTGCTC 400
Qy 181 ATCAACCACATTCGGGCCAGAGAGGCTTGAGGTGGCTCTCAGCACTTTGAGCGGATA 240
Db 401 ATCAACCACATTCGGGCCAGAGAGGCTTGAGGTGGCTCTCAGCACTTTGAGCGGATA 460
Qy 241 AACAGGAAGACCTGTGGGAGAGAGGACAGAGAGAGACCTGTGTGAGG----- 288
Db 461 AACAGGAAGACCTGTGGGAGAGAGGACAGAGAGAGACCTGTGTGAGGATACCCACCT 520
Qy 289 ----- 288
Db 521 GGTGGCCGCTCTCACTTGGGAACCAAGTCAACATGCGCTTCGTGAAGTCTCTTTGTCACT 580
Qy 289 -----GATCCCCAGGAACCTACAGGAGCTATGTCCGCAGGAAATCCCGGCTCATG 339
Db 581 CCAAGAAAGATCCCGAGGAACCTACAGGAGCTATGTCCGCAGGAATTCGGGCTCATG 640
Qy 340 GAAGACCGCAATGCGCGCTAGGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTC 399
Db 641 GAAGACCGCAATGCGCGCTAGGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTC 700
Qy 400 CTGCTGTGTGAGGAGCACTCAAAACCCATGCAGTCCAGCAGCAGCTTCTGGACACAGGC 459
Db 701 CTGCTGTGTGAGGAGCACTCAAAACCCATGCAGTCCAGCAGCAGCTTCTGGACACAGGC 760
Qy 460 CGGGACACGCGAGGACCGTGGGACACAGGCTAGCCCCATCAAGATAGAGACCTCTTT 519
Db 761 CGGGACACGCGAGGACCGTGGGACACAGGCTAGCCCCATCAAGATAGAGACCTCTTT 820
Qy 520 GAGCCAGAGGAGCGCCCGAGCCACCGGCAACCGTGTCATGCAAGGCGCGGACGGG 579
Db 821 GAGCCAGAGGAGCGCCCGAGCCACCGGCAACCGTGTCATGCAAGGCGCGGACGGG 880
Qy 580 ATAGGCAAGTCCATGCTGGCACCAAGTGATGCTGCAGCTGGGCGGAGCGGAAGCTCTTC 639
Db 881 ATAGGCAAGTCCATGCTGGCACCAAGTGATGCTGCAGCTGGGCGGAGCGGAAGCTCTTC 940
Qy 640 CAAGGCAGATTTGATTTATCTCTTACATCAACTGCAAGGAGATGAACAGAGTGCCACG 699
Db 941 CAAGGCAGATTTGATTTATCTCTTCTATCATCAACTGCAAGGAGATGAACAGAGTGCCACG 1000
Qy 700 GAATGCAAGCATGCAAGACCTCATCTTCAGTGTGGCTGAGCCAGCGCGCTCTCCAG 759
Db 1001 GAATGCAAGCATGCAAGACCTCATCTTCAGTGTGGCTGAGCCAGCGCGCTCTCCAG 1060
Qy 760 GAGCTCATCCGAGTTCGAGCGCTCTTTTTCATCATCGACGCTTCGATGAGCTCAAG 819
Db 1061 GAGCTCATCCGAGTTCGAGCGCTCTTTTTCATCATCGACGCTTCGATGAGCTCAAG 1120
Qy 820 CCTTTTTCACGATCCTCAGGGACCTGGTGCTCTGCTGGGAGGAGAAACGGCCCAAG 879
Db 1121 CCTTTTTCACGATCCTCAGGGACCTGGTGCTCTGCTGGGAGGAGAAACGGCCCAAG 1180
Qy 880 GAGCTGCTTTTAAACAGCTTAATTCGAAGAGTGTCTCCTGAGCTATCTTTGCTCATC 939
Db 1181 GAGCTGCTTTTAAACAGCTTAATTCGAAGAGTGTCTCCTGAGCTATCTTTGCTCATC 1240
Qy 940 ACCACACGGCCCGCTTTGGAGAGCTCCACGCTGTCTGAGCAGACCCAGGCAATGTG 999
Db 1241 ACCACACGGCCCGCTTTGGAGAGCTCCACGCTGTCTGAGCAGACCCAGGCAATGTG 1300
Qy 1000 GAGATCTGGCTTCTCTGAGGAGAGAAAGGAATACTTCTCAAGTATTTCCACAAT 1059
Db 1301 GAGATCTGGCTTCTCTGAGGAGAGAAAGGAATACTTCTCAAGTATTTCCACAAT 1360
Qy 1060 GCAGAGAGCGGGCCCAAGTCTTCAATTAAGTGGGACCAACGAGCCTCTCTTCCACCATG 1119
Db 1361 GCAGAGAGCGGGCCCAAGTCTTCAATTAAGTGGGACCAACGAGCCTCTCTTCCACCATG 1420
Qy 1120 TGCTTCGTCCTCGTGTGTGCTGGGTGTGTACCTGCTCCAGCAGCAGCTGGAGGT 1179
Db 1421 TGCTTCGTCCTCGTGTGTGCTGGGTGTGTGTACCTGCTCCAGCAGCAGCTGGAGGT 1480

Qy 1180 GGGGGGCTGTTGAGACAGACGCTCCAGGACCACTGCACTGCTATGCTCTACCTGCTG 1239
Db 1481 GGGGGGCTGTTGAGACAGACGCTCCAGGACCACTGCACTGCTATGCTCTACCTGCTG 1540
Qy 1240 AGTCTGATGCAACCCAAAGCCGGGGGCCCGGCTCCAGCCCCCACCACCAAGAGAGGG 1299
Db 1541 AGTCTGATGCAACCCAAAGCCGGGGGCCCGGCTCCAGCCCCCACCACCAAGAGAGGG 1600
Qy 1300 TTGTGCTCTCTTGGCGCAGATGGCTCTGGAATCAGAAAATCTTATTTGAGGAGCAGAC 1359
Db 1601 TTGTGCTCTCTTGGCGCAGATGGCTCTGGAATCAGAAAATCTTATTTGAGGAGCAGAC 1660
Qy 1360 CTCGGAAGCA CGGCTTAGACGGGGAAGACGCTCTCTGCTCTCTCTCAACATCATCTTC 1419
Db 1661 CTCGGAAGCA CGGCTTAGACGGGGAAGACGCTCTCTGCTCTCTCTCAACATCATCTTC 1720
Qy 1420 CAGAAAGACATCAACTGTGTGAGAGTACTACAGCTTCTATCTCACTTGAGTTTCCAGGAATTC 1479
Db 1721 CAGAAAGACATCAACTGTGTGAGAGTACTACAGCTTCTATCTCACTTGAGTTTCCAGGAATTC 1780
Qy 1480 TTTGCACTATGTAATATATCTTGGACGAGGGGAGGGCGGGCAGGCCCCAGACACAGAC 1539
Db 1781 TTTGCACTATGTAATATATCTTGGACGAGGGGAGGGCGGGCAGGCCCCAGACACAGAC 1840
Qy 1540 GTGACCAAGGCTGTGTGACCGGATACGGCTTTTCTGAAGAGAGCTTCTCTGGCACTCACCAGC 1599
Db 1841 GTGACCAAGGCTGTGTGACCGGATACGGCTTTTCTGAAGAGAGCTTCTCTGGCACTCACCAGC 1900
Qy 1600 CGCTTCTGTTTGGACTCTGTAACGAGGAGACACAGAGGCCACTTGGAGAAAGTCTCTGC 1659
Db 1901 CGCTTCTGTTTGGACTCTGTAACGAGGAGACACAGAGGCCACTTGGAGAAAGTCTCTGC 1960
Qy 1660 TGAAGGCTCGCGGCACATCAAGATGAGCCTGTTGCAAGTGGATCCAAAGCAAGCTCAG 1719
Db 1961 TGAAGGCTCGCGGCACATCAAGATGAGCCTGTTGCAAGTGGATCCAAAGCAAGCTCAG 2020
Qy 1720 AGGCAAGGCTCCACCTTGCAAGAGGCTCTCTGGAGTTCTTCAAGTCTCTTGTACGAGATC 1779
Db 2021 AGGCAAGGCTCCACCTTGCAAGAGGCTCTCTGGAGTTCTTCAAGTCTCTTGTACGAGATC 2080
Qy 1780 CAGGAGAGAGTATTCAGAGAGGCTCTGAGCCACTTCCAGGTGATCTGCTGAGTCAAGAAC 1839
Db 2081 CAGGAGAGAGTATTCAGAGAGGCTCTGAGCCACTTCCAGGTGATCTGCTGAGTCAAGAAC 2140
Qy 1840 ATTGCTCTCAAGATGAGGACATGCTCTCTCTCTGCTGCTGAGCGCTCAGAGAGCGCC 1899
Db 2141 ATTGCTCTCAAGATGAGGACATGCTCTCTCTCTCTGCTGCTGAGCGCTCAGAGAGCGCC 2200
Qy 1900 CAGGTGCTGCACTTGTATGGCGCACCTTACAGCGCGGAGCGGGAAGACCGCGGAGGTGC 1959
Db 2201 CAGGTGCTGCACTTGTATGGCGCACCTTACAGCGCGGAGCGGGAAGACCGCGGAGGTGC 2260
Qy 1960 TCCGAGGAGCGCACGCTGTTGGTGCACTCAGACTCAGACAGAGAGACCGTCTGCTGAGAC 2019
Db 2261 TCCGAGGAGCGCACGCTGTTGGTGCACTCAGACTCAGACAGAGAGACCGTCTGCTGAGAC 2320
Qy 2020 GCCTACAGTGAAATCTGGCAGCGGCTCTGTCACCAATCCAACTCTGATAGAGCTGCT 2079
Db 2321 GCCTACAGTGAAATCTGGCAGCGGCTCTGTCACCAATCCAACTCTGATAGAGCTGCT 2380
Qy 2080 CTGTACCGAAATGCCCTTGGCAGCGCGGGGGTGAAGCTGCTCTCTCAAGGACTCAGACAC 2139
Db 2381 CTGTACCGAAATGCCCTTGGCAGCGCGGGGGTGAAGCTGCTCTCTCAAGGACTCAGACAC 2440
Qy 2140 CCCAACTGCAAACTTCAGAACTTGAAGGCTGAAGGTCGCCATCTCCAGCTCAGCCTGC 2199
Db 2441 CCCAACTGCAAACTTTCAGAACTTGAAGGCTGAAGGTCGCCATCTCCAGCTCAGCCTGC 2500
Qy 2200 GAGGACCTCTCGAGCTCTCATAGCAATAAGAAATTTGACAGGATGGATCTCAGTGGC 2259
Db 2501 GAGGACCTCTCGAGCTCTCATAGCAATAAGAAATTTGACAGGATGGATCTCAGTGGC 2560

disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer.

Claim 1; SEQ ID NO 398; 1185pp; English.

The invention relates to 971 novel human cDNA sequences (ADC39919-ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The invention also relates to nucleic acid sequences over 99% identical with the novel human cDNAs. The invention additionally encompasses expression vectors and host cells comprising a nucleic acid of the invention; the recombinant production of a polypeptide of the invention; an antibody against a polypeptide of the invention; a method of detecting an antigenic polynucleotide or polypeptide of the invention; and methods of identifying a compound which binds to a polypeptide of the invention. The invention further discloses methods of preventing, treating or ameliorating a medical condition, kits comprising polynucleotide probes and/or monoclonal antibodies for carrying out the methods of the invention; methods for the identification of compounds that modulate the expression or activity of the polynucleotide and/or polypeptide; and 767 contig sequences corresponding to the cDNA sequences of the invention (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628-ADC33394). The nucleic acids and polypeptides of the invention are useful in diagnostics, drug screening, forensics, gene mapping, in the identification of mutations responsible for genetic disorders or other traits, for assessing biodiversity, and in producing many other types of data and products dependent on DNA and amino acid sequences. They are also used for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, anaemia, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer. The nucleic acids may also be used as hybridisation probes or primers, and in the recombinant production of a protein. The polypeptides are also useful in generating antibodies, as molecular weight markers, and as food supplements. The present sequence represents a specifically claimed human cDNA sequence of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at: ftp.wipo.int/pub/published_pct_sequences.

Sequence 3306 BP; 718 A; 958 C; 939 G; 691 T; 0 U; 0 Other;

Query Match 78.8%; Score 2449.4; DB 10; Length 3306;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 2644; Conservative 1; Mismatches 1; Indels 174; Gaps 2;

290 ATCCCAGAAACCTACAGGACTATGTCGCCGAGAAATTCGGCTCATGAAGACCCCA 349
498 ATCCCAGAAACCTACAGGACTATGTCGCCGAGAAATTCGGCTCATGAAGACCCCA 557
350 ATGCGCGCTAGGGGAATGTCAACCTCAGCCACCGGTACACCGGCTCTGTGGTGA 409
558 ATGCGCGCTAGGGGAATGTCAACCTCAGCCACCGGTACACCGGCTCTGTGGTGA 617
410 AGGAGCACTCAAAACCCATGCGAGGTTCAGACAGCTTCGACACAGCCGGGACAG 469
618 AGGAGCACTCAAAACCCATGCGAGGTTCAGACAGCTTCGACACAGCCGGGACAG 677
470 CGAGGACCTGGGACACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGAG 529
678 CGAGGACCTGGGACACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGAG 737
530 AGGAGCGCCCGGAGCCACCGCGCATCGTGTATGAAGCGCGGAGGATAGGCAACT 589
738 AGGAGCGCCCGGAGCCACCGCGCATCGTGTATGAAGCGCGGAGGATAGGCAACT 797
590 CCATGTGGGACACACAGGCTAGCTGGCTGGCGGACCGGAGCTCTTCCAGGCGAGAT 649
798 CCATGTGGGACACACAGGCTAGCTGGCTGGCGGACCGGAGCTCTTCCAGGCGAGAT 857
650 TTGATTATCTCTTACATCAACTGCGAGGAGATGAACAGAGTGCACCGAATGACCA 709
858 TTGATTATCTCTTACATCAACTGCGAGGAGATGAACAGAGTGCACCGAATGACCA 917
710 TGCAAGACCTCATCTTCAGCTGTGGCTGAGCCGCGCTCTCCAGGAGCTCATCC 769

2260 AACGGCTTGGATTCCAGGATGATGCTGCTTTGCGAGGCTCGGGCATCCCCAGTGC 2319
2561 AACGGCTTGGATTCCAGGATGATGCTGCTTTGCGAGGCTCGGGCATCCCCAGTGC 2620
2320 AGGCTGCAGATGATTCAGTTGAGGAAGTCTAGCTGGAGTCCGGGGCTTGTGAGAGATG 2379
2621 AGGCTGCAGATGATTCAGTTGAGGAAGTCTAGCTGGAGTCCGGGGCTTGTGAGAGATG 2680
2380 GCTTCTGTGCTCGGACCAACCATCTGTTGAGTTCGAGCTGACAGGAATGCACTG 2439
2681 GCTTCTGTGCTCGGACCAACCATCTGTTGAGTTCGAGCTGACAGGAATGCACTG 2740
2440 GAGGATTTGGCTGAGCTTACTATGCGAGGACTGAGGACCCAGTCTGCGAGATACGG 2499
2741 GAGGATTTGGCTGAGCTTACTATGCGAGGACTGAGGACCCAGTCTGCGAGATACGG 2800
2500 ACTTTGGCTGAGAT-----CTGCCGCTCAGCTGCTGCTGCTGACAGGCTGGCC 2553
2801 ACTTTGGCTGCTGGCTGAGTGTGCGCTCAGAGCCAGGCTTGTGAGAACTTTTAC 2860
2554 TCAACTCTCAGTGTGAACACGAGCTGAGAGAGCTGAGACCTGAGCCCTGAATGAGTGGG 2613
2861 TTCACCTCTGGGATCAACGACCTTGACCGACCTTTACCTGACCAACAGCCCTAGGG 2920
2614 GACCTCGGGGTGCTGCTGCTGTGTGAGGGCTCAGGATCCACGTCGACAGTCCAGACC 2673
2921 GACACAGGTGCTCGACTGCTTTGCAAGCGGTGAGCCATCTGGCTGCAAACTCCGAGTC 2980
2674 CTGCGGTT 2681
2981 CTCTGGTT 2988

RESULT 11

ADC30316
ID ADC30316 standard; cDNA; 3306 BP.

AC ADC30316;

DT 18-DEC-2003 (first entry)

XX Human novel cDNA sequence, SEQ ID NO:398.

DE Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianemic; anticoagulant; thrombolytic; vulnery;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 19; gene; ss.

XX Homo sapiens.

XX WO2003029271-A2.

XX 10-APR-2003.

XX 24-SEP-2002; 2002WO-US030474.

XX 24-SEP-2001; 2001US-0324631P.

XX (HYSE-) HYSEQ INC.

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Wang G;

XX Haley-Vicente D, Drmanac RT;

XX WPI; 2003-371981/35.

XX P-PSDB; ADC31287.

XX New polynucleotide and polypeptide useful for diagnosing, preventing or treating conditions such as neurodegenerative diseases, anemias, platelet

| | | | | | | |
|----------------------------|---|---|------|------|---|------|
| QY | 2930 | TTTACCTGACCAACAGCCCTAGGGGACACAGGTGTCGAGTCTGCTTTGCAAGCGGCTGA | 2989 | 1 | ATGCTACGACCCGACGACGCGCTCTGTGCTCTGTCCACCTACTTGGGAAGAACTC | 60 |
| DB | 2964 | TTTACCTGACCAACAGCCCTAGGGGACACAGGTGTCGAGTCTGCTTTGCAAGCGGCTGA | 3023 | 221 | ATGCTACGAAACCGACGACGAGCGCTCTGTGCTCTGTCCACCTACTTGGGAAGAACTC | 280 |
| QY | 2990 | GCCATCTGCTGCTGCAAACTCCGAGTCTCTGCTGTTATTTGGGATGGACCTGAATAAATGA | 3049 | 61 | GAGGCTCTGGAACTGAAGAAAGTTCAAGTTTATACTGGGGAACCGACGACAGAGCTGGGAGAA | 120 |
| DB | 3024 | GCCATCTGCTGCTGCAAACTCCGAGTCTCTGCTGTTATTTGGGATGGACCTGAATAAATGA | 3083 | 281 | GAGGCTCTGGAACTGAAGAAAGTTCAAGTTTATACTGGGGAACCGACGACAGAGCTGGGAGAA | 340 |
| QY | 3050 | CCACAGTAGTTGGCAGCGCTTCGAGTAACAAAACCTTATTGGACATTTGGCTGCTGA | 3108 | 121 | GGCAAGATCCCTCTGGGGAACGATGGGAGGCGGCTCCCTCGAAATGGCCACAGCTGCTC | 180 |
| DB | 3084 | CCACAGTAGTTGGCAGCGCTTCGAGTAACAAAACCTTATTGGACATTTGGCTGCTGA | 3142 | 341 | GGCAAGATCCCTCTGGGGAACGATGGGAGGCGGCTCCCTCGAAATGGCCACAGCTGCTC | 400 |
| RESULT 12 | | | | 181 | ATACCCCACTTCGGGCGCAGAGGAGGCTTGAGGTGTGCTCTCAGCACTTTTGAAGCGGATA | 240 |
| ADP47729 | | | | 401 | ATCACCCACTTCGGGCGCAGAGGAGGCTTGAGGTGTGCTCTCAGCACTTTTGAAGCGGATA | 460 |
| XX | ADP47729; | | | 241 | AACAGGAAGGACCTGTGGGAGAGGAGACAGAGAGGACCTGGTGAGG----- | 288 |
| AC | | | | 461 | AACAGGAAGGACCTGTGGGAGAGGAGACAGAGAGGACCTGGTGAGGATACCCACCT | 520 |
| XX | 12-AUG-2004 (first entry) | | | 289 | ----- | 288 |
| DT | | | | 521 | GGTGGCCGCTCCTCACTTGGGAAACAGTCAACATGCCTTCTGGAAGTCTCTCTGTGCACT | 580 |
| XX | Human Monarch-1 (CATERPILLER 19.3) isoform IV cDNA. | | | 289 | -----GATCCCCCAGGAAACCTACAGGACTATGTCCGACGAAATTCGGGCTCATG | 339 |
| XX | Monarch-1; CATERPILLER 11.2; caspase recruitment domain; | | | 581 | CCAGAAAGATCCCCAGGAAACCTACAGGACTATGTCCGACGAAATTCGGGCTCATG | 640 |
| KW | CARD, transcription enhancer, R-binding, pyrin, lots of leucine repeat; | | | 340 | GAAGACCGCAATCGCGGCTTAGGGGAAATGTCTCAACTCAGCCACCGGTACACCCGGCTC | 399 |
| KW | CATERPILLER 11.3; CATERPILLER 16.1; CATERPILLER 16.2; CIAS1; | | | 641 | GAAGACCGCAATCGCGGCTTAGGGGAAATGTCTCAACTCAGCCACCGGTACACCCGGCTC | 700 |
| KW | cold-induced autoinflammatory syndrome 1; antiinflammatory; cytostatic; | | | 400 | CTGCTGGTGAAGGAGCACTCAAAACCCCATGCAGGTCCAGAGAGGCTTCTGGACACAGGC | 459 |
| KW | inflammatory disease; cancer; gene therapy; human; gene; ss; purine; | | | 701 | CTGCTGGTGAAGGAGCACTCAAAACCCCATGCAGGTCCAGAGAGGCTTCTGGACACAGGC | 760 |
| KW | CATERPILLER 19.3; chromosome 19q13; multiple sclerosis; isoform IV. | | | 460 | CGGGACACGCGAGGACCGTGGGACACAGGCTAGCCCCATCAGATAGAGAGACCTCTTT | 519 |
| OS | Homo sapiens. | | | 761 | CGGGACACGCGAGGACCGTGGGACACAGGCTAGCCCCATCAGATAGAGAGACCTCTTT | 820 |
| XX | WO2004034093-A2. | | | 520 | GAGCCAGAGGAGGCGGCCCGGACACCGGCAACCGGCAACCGTGGTCAATCAAGGCGCGCAGGG | 579 |
| XX | 22-APR-2004. | | | 821 | GAGCCAGAGGAGGCGGCCCGGACACCGGCAACCGGCAACCGTGGTCAATCAAGGCGCGCAGGG | 880 |
| XX | 30-APR-2003; 2003WO-US013562. | | | 580 | ATAGGCAAGTCCATGCTGGGACACAGAGGTGATGCTGGACTGGCGGAGCGGGAAGCTCTTC | 639 |
| XX | 30-APR-2002; 2002US-0376626P. | | | 881 | ATAGGCAAGTCCATGCTGGGACACAGAGGTGATGCTGGACTGGCGGAGCGGGAAGCTCTTC | 940 |
| XX | (UYNC-) UNIV NORTH CAROLINA. | | | 640 | CAAGGCAAGTTTGAATATCTCTTCTACATCAACTGCAGGAGATGAACCCAGAGTGCCACG | 699 |
| XX | Ting JY, Linhoff MW, Harton JA, Williams KL, Lich J, O'connor W; | | | 941 | CAAGGCAAGTTTGAATATCTCTTCTACATCAACTGCAGGAGATGAACCCAGAGTGCCACG | 1000 |
| PI | Moore CB, Davis B, Brickey J, Conti B, Zhang J, Zhu X; | | | 700 | GAATGACAGATCAAGACCTCATCTTTCAGTGTGGCTTGAGCCAGCGCGCTCTCCAG | 759 |
| XX | WPI; 2004-348215/32. | | | 1001 | GAATGACAGATCAAGACCTCATCTTTCAGTGTGGCTTGAGCCAGCGCGCTCTCCAG | 1060 |
| DR | P-PSDB; ADP47730. | | | 760 | GAGCTATCCGAGTTCCGAGCGCTCTCTTTTCATCATCGAGGCTTCGATGAGTCAAG | 819 |
| XX | New nucleic acid encoding Monarch-1, CATERPILLER 11.2, CATERPILLER 11.3, | | | 1061 | GAGCTATCCGAGTTCCGAGCGCTCTCTTTTCATCATCGAGGCTTCGATGAGTCAAG | 1120 |
| PT | CATERPILLER 16.1, CATERPILLER 16.2 or CIAS1 polypeptide, useful in | | | 820 | CCCTCTTTCCACGATCTCAGGAGACCTGGTGGCTCTGCTGGGAGAGAAACCGGCCACG | 879 |
| PT | preparing a composition for treating inflammatory disease or cancer. | | | 1121 | CCCTCTTTCCACGATCTCAGGAGACCTGGTGGCTCTGCTGGGAGAGAAACCGGCCACG | 1180 |
| XX | Claim 4; SEQ ID NO 7; 205pp; English. | | | 880 | GAGCTGCTTTTAAACAGCTTAATCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATC | 939 |
| CC | The invention relates to a novel isolated nucleic acid encoding a Monarch | | | 1181 | GAGCTGCTTTTAAACAGCTTAATCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATC | 1240 |
| CC | -1, CATERPILLER (CARD [caspase recruitment domain], transcription | | | 940 | ACCACAGGCGCCCGGCTTTGGAGAACTCCACCGTCTGCTGGAGACCCACGAGCATGTG | 999 |
| CC | enhancer, R(purine)-binding, pyrin, lots of leucine repeats) 11.2, | | | 1241 | ACCACAGGCGCCCGGCTTTGGAGAACTCCACCGTCTGCTGGAGACCCACGAGCATGTG | 1300 |
| CC | CATERPILLER 11.3, CATERPILLER 16.1, CATERPILLER 16.2 or CIAS1 (cold- | | | | | |
| CC | induced autoinflammatory syndrome 1) polypeptide comprising the amino | | | | | |
| CC | acids encoded by exon 6 and lacking the amino acids encoded by exon 4 or | | | | | |
| CC | its fragment. The nucleic acid of the invention demonstrates | | | | | |
| CC | antiinflammatory and cytostatic activities and may be useful in preparing | | | | | |
| CC | a composition for treating an inflammatory disease or cancer, possibly | | | | | |
| CC | via gene therapy. The current sequence is that of the human Monarch-1 | | | | | |
| CC | (CATERPILLER 19.3) isoform IV cDNA of the invention which is located on | | | | | |
| CC | chromosome 19q13, in the multiple sclerosis susceptibility region. | | | | | |
| XX | Sequence 3221 BP; 749 A; 891 C; 936 G; 645 T; 0 U; 0 Other; | | | | | |
| SQL | | | | | | |
| Query Match | 77.8%; Score 2417; DB 12; Length 3221; | | | | | |
| Best Local Similarity | 96.9%; Pred. No. 0; | | | | | |
| Matches 2508; Conservative | 0; Mismatches 0; Indels 81; Gaps 1; | | | | | |

The invention relates to a novel isolated nucleic acid encoding a Monarch
CC -1, CATERPILLER (CARD [caspase recruitment domain], transcription
CC enhancer, R(purine)-binding, pyrin, lots of leucine repeats) 11.2,
CC CATERPILLER 11.3, CATERPILLER 16.1, CATERPILLER 16.2 or CIAS1 (gold-
CC induced autoinflammatory syndrome 1) polypeptide comprising the amino
CC acids encoded by exon 6 and lacking the amino acids encoded by exon 4 or
CC its fragment. The nucleic acid of the invention demonstrates
CC antiinflammatory and cytostatic activities and may be useful in preparing
CC a composition for treating an inflammatory disease or cancer, possibly
CC via gene therapy. The current sequence is that of the murine Monarch-1
CC (CATERPILLER 19.3) predicted DNA of the invention.

XX
SQ Sequence 3102 BP; 750 A; 786 C; 872 G; 694 T; 0 U; 0 Other;

| | | | | |
|-----------------------|-----------------|-----------------|------------|--------------|
| Query Match | 60.8%; | Score 1864.8; | DB 12; | Length 3102; |
| Best Local Similarity | 75.7%; | Pred. No. 0; | | |
| Matches 2367; | Conservative 0; | Mismatches 717; | Indels 42; | Gaps 3; |

| | | | |
|----|-----|---|-----|
| Qy | 1 | ATGCTACGAACCGCAGCAGGAGCGGCTCTGTGCGCTGTCCACCTACTTGGNAGAACTC | 60 |
| Db | 1 | ATGTTGCGCTACAGCCAGGATGGGCTCTATCGACTGTCTACCTACCTTGGGAAGAACTC | 60 |
| Qy | 61 | GAGCTGTGGAACTGAGAGATTCAAGTTATACCTGGGACCGCAGACAGACTGGGAGAA | 120 |
| Db | 61 | GAGCTGTGGGAACTGAGAGAAATTCAAATTTATCTCTGGGGATTGAGAGGACCTGAGCCAG | 120 |
| Qy | 121 | GGCAAGATCCCTGGGGAAGCATGGAGAGCCGGTCCCTTGGAAATGGCCCGAGCTGCTC | 180 |
| Db | 121 | GACAAATTCCTGGGACGAATGGAGAGCGTGGTCTCTGGGAAATGGCTCAGCTGATG | 180 |
| Qy | 181 | ATCACCCATCTGGGCGCAGAGAGCCCTGAGGTTGGCTCTCAGCACCTTTGAGCGGATA | 240 |
| Db | 181 | GTGGCCCATATGGGACAAAGGAGGCTTGGCTTCTGGGCTCTCAGCACCTTTTCAAGAT | 240 |
| Qy | 241 | AACAGGAGGACCTGTGGGAGAGGACAGAGAGGACCTGTGTGAGG----- | 288 |
| Db | 241 | CACAGGAGGACCTGTGGGAGGAGGACAGGAGAGACCTGTGTGAGGGGTAAAGAGGGC | 300 |
| Qy | 289 | -----GATCCCCAGGAAACCTACAGGACATATGTCCGAGGAATTCGGCTCATGGAA | 342 |
| Db | 301 | AAGGGAGATCTACAGCAACCTTACAAAGACTATGTCCGAGGAATTTCCAGCTAATGGAA | 360 |
| Qy | 343 | GACCCCATGTGGCGCTAGGGGATGTGTCAACCTTCAGCCACCGGTACACCCGGCTCCTG | 402 |
| Db | 361 | GACCGCAATGTCAAGATTAGGCGAATGTGTGAACCTGAGCAATCGTTACACTCGGCTTCTC | 420 |
| Qy | 403 | CTGGTGAAGGAGCACTCAAAACCCATGCAAGTCCAGAGAGAGCTTCTGGAACAGGCGG | 462 |
| Db | 421 | CTAGTAAAGAACACTCAAACTCTATCTGGACACAGCAGAGAAATTTGTAGATGTAGATGG | 480 |
| Qy | 463 | GGACGCGAGGACCGTGGGACACAGGCTAGCCCATCAGATAGACACCTCTTTGAG | 522 |
| Db | 481 | GAACTGTGAGAACCAAGGCTCACCAGACTAGTCTCTATCCAAATGGAGACCTCTTTGAG | 540 |
| Qy | 523 | CCAGAGAGAGGCGCCCGAGCACCGCGCACCGTGTGTATGCAAGGGCGCGCAGGATA | 582 |
| Db | 541 | CCAGAGAGAGAGCGCCCGAGGCCACACACAGTGTATACAAAGGGCGCAGCGGGATG | 600 |
| Qy | 583 | GGCAAGTCCATGTGCGACACAAAGGTGATGTGTGACTGGCGGAGCGGAAGCTTCCAA | 642 |
| Db | 601 | GGGAACTCATGTGCGCCACAAAGGTGATGTGTGACTGGCGGCGATGGGAGGCTTCCAA | 660 |
| Qy | 643 | GGCAGATTTGATTTATCTCTTACATCAACTGCAGGAGATGAACACAGAGTCCACAGGAA | 702 |
| Db | 661 | GGCGGTTGATTTATGTCTTATATACGTGAGGAGTTGATAGAGCCACACCCAG | 720 |
| Qy | 703 | TGCAGATGCAAGACCTCATCTTTCACTGTGTGCTGAGCCCGAGCGGCTCTCCAGGAG | 762 |
| Db | 721 | TGCAGTGTACAAAGACCTATCTCCAGCTGTGTGGCGGAGCGTGGTATATCCCTCGAAGAC | 780 |
| Qy | 763 | CTCATCCGATTTCCGAGGCGCTCTTTTCATCATGAGGCTTCGATGAGCTCAAGCT | 822 |
| Db | 781 | CTCATGAGGCTCTGACCGCTCTCTTATTCATCATTTGATGGCTTCGATTAATCCATCTCT | 840 |

| | | | |
|----|------|---|------|
| Qy | 823 | TCCTTCCACGATCCTCAGGACACCTGTGTGCTCTGTGGGAGGAGAAACGCCCAACGAG | 882 |
| Db | 841 | TCCTTCCATGATCTCAGGCTCCTGTGTGCTCTGTGGGAGGAGAAACCACTACTGAA | 900 |
| Qy | 883 | CTGCTTCTTAACAGCTTAAATTCGGAAGAGCTGCTCCTCAGCTATCTTTGCTCATCACC | 942 |
| Db | 901 | GTCTCTCTCGGAAGCTGATTCGGAGGTGCTTCTGCCAGGCTCTCTGCTCATCACC | 960 |
| Qy | 943 | ACAGGCGCACCGCTTTGGAGAGCTCCAAGCTGCTGTGGAGCACCCAGGCAATGGAG | 1002 |
| Db | 961 | ACAGGACCTGTGCACTGGAGAGCTGCACGGCTTGTAGAACACCCAGGACGCTGAG | 1020 |
| Qy | 1003 | ATCTGTGGCTTCTCTGAGGAGAAAGGAATACTTCTACAGTATTTTCCAAATGCA | 1062 |
| Db | 1021 | ATCTGTGGCTTCTCGAGGAGCTAGGAGGAATTTCTACAGATATTTCCAAACACT | 1080 |
| Qy | 1063 | GAGCAGGCGGCGCCAAAGTCTTCAATTACGTGAGGAGCAACAGAGCTCTTTCACCATGTC | 1122 |
| Db | 1081 | GGACAAGCAAGCCGGGTGTTAAGCTTCTTGAATGACTATGAGCCCTCTTTTACCATGTG | 1140 |
| Qy | 1123 | TTGCTCCCTGCTGTGCTGGGTGTGTACTGCTTCCAGCAGCAGCTCTTTCAGGAGGGTGG | 1182 |
| Db | 1141 | TTTGTTCCTCATGGTGTCTGGGTGTGTGACCTTGTGACCTTAAAGCAGCAGCTGGAAGTGG | 1200 |
| Qy | 1183 | GGGCTGTTGAGACAGACGCTCCAGGACCACTGCAAGTGTACATGCTCTACTGCTGAGT | 1242 |
| Db | 1201 | GAGCTTTTAAAGCAAAACACTTAGGACCACTAGGAGCTGTTTATATGTTCTACTTCTGAGC | 1260 |
| Qy | 1243 | CTGATGCAACCCCAAGCCGGGCGCCCTCCAGCCGCCCAACCCAGAGAGGGTGG | 1302 |
| Db | 1261 | CTGATGCAACCCCAAGCCAGGAGCTCAACCTTCAAGTCCCGAGCCCAACAGAGGCGCTG | 1320 |
| Qy | 1303 | TGCTCTCTGGGCGCAGATGGGCTCTGGAATCAGAAATCTTATTTGAGGACGACGACCTC | 1362 |
| Db | 1321 | GTCTCTCTGGCTGAGAGGGCTCTGGAATCAGAGATTTCTATTTGATGAACAGGATCTT | 1380 |
| Qy | 1363 | CGGAAGCAACGGCTTGAACGGGAAACGCTCTCTGCTTCTTCAACATGAATCTTCCAG | 1422 |
| Db | 1381 | GGGAACACGGCTAGATGGAGCAGATGTGTCTCTTCTTCAACGTAACATATTTCCAG | 1440 |
| Qy | 1423 | AAGGACATCACTGTGAGAGGTACTACAGCTTCTTCACTTCTGAGTTTCCAGGAATTTCTTT | 1482 |
| Db | 1441 | AAGGGTATCAATGTGAGAAATTTCTACAGCTTCTTCACTTCTGAGTTTCCAGGAATTTCTT | 1500 |
| Qy | 1483 | GCAGCTATGTACTATATCTCTGACAGGGGGAGGGGGGAGGCCAGCCAGACGACGCTG | 1542 |
| Db | 1501 | GCAGCCATGTACTGTGCACTGAATGGCAGAG-----AGGCGGTG | 1539 |
| Qy | 1543 | ACCAGGCTGTGACCGAGTACGGCTTCTTGAAGGAGCTTCTTGGCACTTCCAGCCGC | 1602 |
| Db | 1540 | AGGAGAGCGCTGGCTGAGTATGGTTTTTTCGGAAGGAACTTTCTTGGCCCTTCCGGTCCAC | 1599 |
| Qy | 1603 | TTCTGTTTGGACTCTCTGAAACGAGGAGACAGGAGCCACTTGGAGAGAGTCTCTGCTGG | 1662 |
| Db | 1600 | TTTCTGTTTGGCTCTCTCAACGAGAGATGAGTCTACTCTTGAAGGAATCTCGGCTGG | 1659 |
| Qy | 1663 | AAGGTCTTCGCGCACATCAAGATGAGCCTGTTGCACTGATTCCTCAAGCAAGCTCAGAGC | 1722 |
| Db | 1660 | AGCATCTCCCTCAGGTGAGGAGGAAGTGTGGCATGGATCCAAACAGGCTGGGAGT | 1719 |
| Qy | 1723 | GACGGCTCCACCTCGAGCAGGGCTCTTGGAGTCTTCTAGCTGCTTGTACGAGATCCAG | 1782 |
| Db | 1720 | GAGGGCTCCACCTTGCAGCATGGCTCTCTGAGACTACTCAGCTGCTTGTATGAGGCTCCAG | 1779 |
| Qy | 1783 | GAGGAGGAGTTTATCCAGCAGGCGCTTGAGCCAATTCCAGGTGATCGTGTGTCAGCAAT | 1842 |
| Db | 1780 | GAGGAGGACTTCATCCAGCAGGCGCTGAGCCAATTTCAGTGGTGTGTAGTCAGAGCATC | 1839 |
| Qy | 1843 | GCCTCCAGAGTGGAGCAGTGGTCTCTGCTGCTGAGCGCTGAGAGCGCCAG | 1902 |
| Db | 1840 | TCAACAAAGATGGAGCAGTGGTCTGCTGCTTGTGCGGAGTATTTGCAAGATGACAGAA | 1899 |

XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX
PS Claim 20; Page 98-99; 332pp; English.
XX
XX The invention describes an isolated polypeptide (I) comprising any of 27
CC 118-961 residue amino acid sequences, given in the specification, a
CC mature form of them, a sequence that is at least 95 % identical to them,
CC or a sequence having one or more conservative substitutions in them. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease selected from a pathology
CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
CC and antibodies are useful in treating or preventing NOVX-associated
CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
CC associated cachexia, and other wasting disorders associated with chronic
CC diseases. The nucleic acids and polypeptides may also be used as targets
CC for the identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods. The nucleic acids are further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, and pharmacogenomics. The polypeptides are also useful as
CC vaccines. This sequence encodes a novel human G-protein coupled receptor
CC related protein NOV
XX
SQ Sequence 1800 BP; 366 A; 502 C; 565 G; 367 T; 0 U; 0 Other;

Query Match 54.9%; Score 1705.4; DB 8; Length 1800;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1748; Conservative 0; Mismatches 1; Indels 12; Gaps 3;

QY 1102 GAGCCTCTCTTCCACATGTGCTTCCCTCCCTGTGCTGGTGTGTACCTGCTC 1161
DB 1 GAGCCTCTCTTCCACATGTGCTTCCCTCCCTGTGCTGGTGTGTACCTGCTC 60

QY 1162 CAGCAGCAGCTGAGGCTGGGGGCTGTTGAGACAGAGCTCCAGACACCACTGCAGTG 1221
DB 61 CAGCAGCAGCTGAGGCTGGGGGCTGTTGAGACAGAGCTCCAGACACCACTGCAGTG 120

QY 1222 TACATGCTCTACTGCTGAGTCTGATGCAACCCAGCCGGGGCCCCCGCTCCAGCCC 1281
DB 121 TACATGCTCTACTGCTGAGTCTGATGCAACCCAGCCGGGGCCCCCGCTCCAGCCC 180

QY 1282 CCACCCAAACAGAGAGGCTTGTGCTTGGCGGAGATGGGCTCTGGAATCAGAAAATC 1341
DB 181 CCACCCAAACAGAGAGGCTTGTCTCCCTTGGCGGAGATGGGCTCTGGAATCAGAAAATC 240

QY 1342 CTATTTTGAAGAGCAGGACCTCCGGAAGCAGCGCTAGACGGGGAAGAGCTCTGCTCCTTC 1401
DB 241 CTATTTTGAAGAGCAGGACCTCCGGAAGCAGCGCTAGACGGGGAAGAGCTCTGCTCCTTC 300

QY 1402 CTCAACATGAACATCTTCCAGAGGACATCACTGTGAGAGGTACTACAGCTTTCATCCAC 1461
DB 301 CTCAACATGAACATCTTCCAGAGGACATCACTGTGAGAGGTACTACAGCTTTCATCCAC 360

QY 1462 TTGAGTTTCCAGAAATTTCTTTGAGCTATGATATATCTGGAAGAGGGGAGGCGGG 1521
DB 361 TTGAGTTTCCAGAAATTTCTTTGAGCTATGATATATCTGGAAGAGGGGAGGCGGG 420

QY 1522 GCAGGCCCAGACCAAGGACGTGAACAGGCTTGTGACCGATACCGCTTTCTGAAAGGAGC 1581
DB 421 GCAGGCCCAGACCAAGGACGTGAACAGGCTTGTGACCGATACCGCTTTCTGAAAGGAGC 480

QY 1582 TTCTTGGAATCTACAGCCGCTTCTGTTGGACTCTGAAACAGAGAGACCAAGAGCCAC 1641
DB 481 TTCTTGGAATCTACAGCCGCTTCTGTTGGACTCTGAAACAGAGAGACCAAGAGCCAC 540

QY 1542 CTGGAGAGAGTCTCTGCTGGAAGTCTCGCGCACATCAAGATGACACTGTTCAGTGG 1701
DB 541 CTGGAGAGAGTCTCTGCTGGAAGTCTCGCGCACATCAAGATGACACTGTTCAGTGG 600

QY 1702 ATCCAAAGCAAAAGCTCAGACGACGCTCCACCTGAGCAGGCTCTTGGAGTCTTC 1761
DB 601 ATCCAAAGCAAAAGCTCAGACGACGCTCCACCTGAGCAGGCTCTTGGAGTCTTC 660

QY 1762 AGCTGCTTACGATCCAGAGGAGGAGTCTTATCCAGCAGGCTCTGAGCCACTTCCAG 1821
DB 661 AGCTGCTTACGATCCAGAGGAGGAGTCTTATCCAGCAGGCTCTGAGCCACTTCCAG 720

QY 1822 GTGATCGTGTGTCAGCAAACTTGCCTCCAGATGAGACATGCTCTCTGTTGCTG 1881
DB 721 GTGATCGTGTGTCAGCAAACTTGCCTCCAGATGAGACATGCTCTCTGTTGCTG 780

QY 1882 AAGCGCTGAGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1941
DB 781 AAGCGCTGAGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840

QY 1942 GAAGACGCGGAGTGTCTCCGAGGAGCGCACGCTGTTGTTGCTGAGCTCAGACAGAG 2001
DB 841 GAAGACGCGGAGTGTCTCCGAGGAGCGCACGCTGTTGTTGCTGAGCTCAGACAGAG 900

QY 2002 AGGACCGTCTGCTGAGCGCTTACAGTGAACATCTGACGAGCGGCTGTGACCAATCCA 2061
DB 901 AGGACCGTCTGCTGAGCGCTTACAGTGAACATCTGACGAGCGGCTGTGACCAATCCA 960

QY 2062 AACCTGATAGAGTGTCTCTGCTGACGAAATGCTGCGGACGCGGGGGTGAAGCTGCTC 2121
DB 961 AACCTGATAGAGTGTCTCTGCTGACGAAATGCTGCGGACGCGGGGGTGAAGCTGCTC 1020

QY 2122 TGTCAAGGACTCAGACACCCCACTGCAAACTTTCAGAACTT---GAGGCTCAAGAGTGC 2178
DB 1021 TGTCAAGGACTCAGACACCCCACTGCAAACTTTCAGAACTTTCAGAGGCTCAAGAGTGC 1080

QY 2179 CGCATCTCAGCTCAGCTGCGAGGACCTCTCTGAGCTCTCATAGCAATTAAGAAATTTG 2238
DB 1081 CGCATCTCAGCTCAGCTGCGAGGACCTCTCTGAGCTCTCATAGCAATTAAGAAATTTG 1140

QY 2239 ACAAGAGTGGATCTCAGTGGCAACCGGCTTGGATTTCCAGGATGATGCTGCTTTCGAG 2298
DB 1141 ACAAGAGTGGATCTCAGTGGCAACCGGCTTGGATTTCCAGGATGATGCTGCTTTCGAG 1200

QY 2299 GGCCTCGGCAATCCCACTGAGGCTCAGATGATTCAGTTCAGGAGTGTGAGCTGAGAG 2358
DB 1201 GGCCTCGGCAATCCCACTGAGGCTCAGATGATTCAGTTCAGGAGTGTGAGCTGAGAG 1260

QY 2359 TCCGGGGCTTGTGAGGAGATGGCTTCTGCTCGGCAACCAACCCACATCTGTTGAGTTG 2418
DB 1261 TCCGGGGCTTGTGAGGAGATGGCTTCTGCTCGGCAACCAACCCACATCTGTTGAGTTG 1320

QY 2418 GACTCAGAGAAATGCACTGGAGATTTGGGCTGAGTTCATGCGAGGACTGAGG 2478
DB 1321 GACTCAGAGAAATGCACTGGAGATTTGGGCTGAGTTCATGCGAGGACTGAGG 1380

QY 2479 CACCCAGTCTGAGACTTACGAGCTTTGT---GGCTGAAGATCTGCGGCTCTCACTGCT 2532
DB 1381 CACCCAGTCTGAGACTTACGAGCTTTGTGGTGGAGCTGGAAGTCTGCGGCTCTCACTGCT 1440

QY 2532 GCTGCTGTCAGAGGCTGGCTCAACTCTCAGTGTGAACAGAGCTGAGAGCTGGAC 2592
DB 1441 GCTGCTGTCAGAGGCTGGCTCAACTCTCAGTGTGAACAGAGCTGAGAGCTGGAC 1500

QY 2593 CTGAGCTGAGTGTGGGAGCTCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2652
DB 1501 CTGAGCTGAGTGTGGGAGCTCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560

QY 2653 CCCAGCTGCAAGCTCCAGACCTGCTG---GGTGTGGGCACTCTGCGGCTGCGGCTCTGCGGCC 2709
DB 1561 CCCAGCTGCAAGCTCCAGACCTGCTGCGGAGGTTGGGCACTCTGCGGCTGCGGCTCTGCGGCC 1620

QY 2710 TGTGAGGCTCTTTCTGTTGGTCTCCAGGCCAACCAACACCTCCGGGAGCTGAGTTCAGT 2769

Db 1621 TGTGAGGCTCTTTCTGTGGTCTCCAGCCCAACACCACTCCGGGAGCTGGACTTGAGT 1680
Qy 2770 TTCAACGACCTGGGAGACTGGGGCTGTGGTTGCTGAGGCGCTGCAACATCCCGCC 2829
Db 1681 TTCAACGACCTGGGAGACTGGGGCTGTGGTTGCTGAGGCGCTGCAACATCCCGCC 1740
Qy 2830 TGAGACTCCAGAACTGTGG 2850
Db 1741 TGCAGACTCCAGAACTGTGG 1761
RESULT 15
ID ABS78719 standard; cDNA; 2158 BP.
XX ABS78719;
XX ABS78719;
XX 16-DEC-2002 (first entry)
XX Human cDNA encoding NAAP9, from INCYTE no.429930CB1.
XX Human; ss; gene; nucleic acid associated protein; NAAP; cancer;
KW cell proliferative disease; cancer; atherosclerosis; hepatitis;
KW neurological disorder; Parkinson's disease; Alzheimer's disease; stroke;
KW epilepsy; developmental disorder; renal tubular acidosis; anaemia;
KW glaucoma; hypothyroidism; autoimmune disorder; AIDS;
KW inflammatory disorder; acquired immunodeficiency syndrome; allergy;
KW atopic dermatitis; arthritis; bacterial infection; viral infection;
KW parasitic infection; protozoal infection; fungal infection.
XX Homo sapiens.
XX WO200272630-A2.
XX 19-SEP-2002.
XX 07-FEB-2002; 2002WO-US003844.
XX 09-FEB-2001; 2001US-0268118P.
XX 21-FEB-2001; 2001US-0270963P.
XX 22-FEB-2001; 2001US-0270858P.
XX 23-FEB-2001; 2001US-0271194P.
XX 07-MAR-2001; 2001US-0274071P.
XX 12-APR-2001; 2001US-0283496P.
XX 09-NOV-2001; 2001US-0344650P.
XX (INCY-) INCYTE GENOMICS INC.
XX Thornton M, Hafalia AJA, Lu DAM, Arvizu C, Swarnakar A, Lu Y;
PI Warren BA, Baughn MR, Tang YT, Lee EA, Yao MG, Ramkumar J, Khan FA;
PI Gandhi AR, Ding L, Yue H, Gietzen KJ, Wallia NK, Thangavelu K;
PI Elliott VS, Marquis JP;
XX WPI; 2002-723320/78.
DR P-PSDB; ABG97475.
XX New human nucleic acid-associated proteins (NAAP), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis,
PT infections.
XX Claim 5; Page 160; 162pp; English.
XX The invention relates to an isolated polypeptide comprising one of 10
CC human nucleic acid associated protein (NAAP1-10), or a biologically
CC active or immunogenic fragment of the polypeptide, and their encoding
CC nucleic acid. Also included are a recombinant polynucleotide comprising a
CC promoter sequence operably linked to the polynucleotide, a cell
CC transformed with the recombinant polynucleotide, a transgenic organism
CC comprising the recombinant polynucleotide, an anti-NAAP antibody,
CC screening for a compound that is effective as an ant/agonist or modulator
CC of NAAP, generating an expression profile of a sample containing the

CC polynucleotides and an array comprising different nucleotide molecules
CC affixed on a solid substrate, nucleotide molecule comprises a first
CC oligonucleotide or polynucleotide sequence specifically hybridizable with
CC at least 30 contiguous nucleotides of the target (NAAP) polynucleotide.
CC The polypeptides and polynucleotides are useful in diagnosing, treating
CC and preventing diseases or conditions associated with the decreased
CC expression or overexpression of NAAP, such as cell proliferative diseases
CC (e.g. cancer, atherosclerosis, hepatitis), neurological disorders
CC (Parkinson's disease, Alzheimer's disease, stroke, epilepsy),
CC developmental disorders (renal tubular acidosis, anaemia, glaucoma,
CC hypothyroidism), autoimmune/inflammatory disorders (AIDS (acquired
CC immunodeficiency syndrome), allergies, atopic dermatitis, arthritis) and
CC infections (e.g. bacterial, viral, parasitic, protozoal, fungal) and many
CC other diseases and disorders listed in the specification. These are also
CC useful in assessing the effects of exogenous compounds on the expression
CC of nucleic acid and amino acid sequences of NAAP. The NAAP or its
CC fragments are useful in screening compounds for effectiveness as agonist
CC or antagonist of the polypeptides, or in altering the expression of the
CC target polynucleotide and compounds that specifically bind to or modulate
CC the activity of the polypeptide. The microarray is useful in monitoring
CC or measuring protein-protein interactions, drug-target interactions, and
CC gene expression profiles. The present sequence encodes an NAAP protein
XX
SQ Sequence 2158 BP; 489 A; 578 C; 639 G; 452 T; 0 U; 0 Other;
Query Match 54.1%; Score 1681.8; DB 6; Length 2158;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1696; Conservative 0; Mismatches 2; Indels 3; Gaps 1;
Qy 1408 ATGAACATCTTCCAGAAGGACATCACTGTGAGAGGTACTACAGCTTCACCTTGA 1467
Db 1 ATGAACATCTTCCAGAAGGACATCACTGTGAGAGGTACTACAGCTTCACCTTGA 60
Qy 1468 TTCAGGAATCTTTCAGACTATGTACTATATCTTGACCGAGTACGCGTTTCTGAAAGGAGCTTCCTG 1527
Db 61 TTCAGGAATCTTTCAGACTATGTACTATATCTTGACCGAGTACGCGTTTCTGAAAGGAGCTTCCTG 120
Qy 1528 CCAGACCGAGACGCTGACAGGCTGTTGACCGAGTACGCGTTTCTGAAAGGAGCTTCCTG 1587
Db 121 CCAGACCGAGACGCTGACAGGCTGTTGACCGAGTACGCGTTTCTGAAAGGAGCTTCCTG 180
Qy 1588 GCACTACAGCGCGCTTCCTGTTTGGACTCTCTGAAAGGAGACGAGAGCCACCTGGAG 1647
Db 181 GCACTACAGCGCGCTTCCTGTTTGGACTCTCTGAAAGGAGACGAGAGCCACCTGGAG 240
Qy 1648 AAGAGTCTCTGCTGGAAGGTCTCGCCGACATCAAGATGACCTGTTGCACTGGATCCAA 1707
Db 241 AAGAGTCTCTGCTGGAAGGTCTCGCCGACATCAAGATGACCTGTTGCACTGGATCCAA 300
Qy 1708 AGCAAGAGCTCAGAGCGACGCGCTCCACCCTGCAGCAGGCGCTCCTTGGAGTTCTTCAGCTGC 1767
Db 301 AGCAAGAGCTCAGAGCGACGCGCTCCACCCTGCAGCAGGCGCTCCTTGGAGTTCTTCAGCTGC 360
Qy 1768 TTGTAAGAGATCCAGAGGAGGAGTTTATCCAGAGCGCTTCAGAGCCACTTCAGAGTATC 1827
Db 361 TTGTAAGAGATCCAGAGGAGGAGTTTATCCAGAGCGCTTCAGAGCCACTTCAGAGTATC 420
Qy 1828 GTGCTCAGCAACATTCGCTCCAGAGTGGAGCACATGCTCTCCTGTTCTCTCTGAAGCGC 1887
Db 421 GTGCTCAGCAACATTCGCTCCAGAGTGGAGCACATGCTCTCCTGTTCTCTCTGAAGCGC 480
Qy 1888 TGCAGAGCGCGCCAGGCTGCTGCATCTTGTATGGCGGCCACCTTACAGCGCGGACGGGGAAGAC 1947
Db 481 TGCAGAGCGCGCCAGGCTGCTGCATCTTGTATGGCGGCCACCTTACAGCGCGGAGGGAAGAC 540
Qy 1948 CGCGAGGCTGCTCCGAGAGGAGCGCAACGCTGTTGGTGCAGCTTCAGACGAGAGAGACC 2007
Db 541 CGCGAGGCTGCTCCGAGAGGAGCGCAACGCTGTTGGTGCAGCT---ACCAGAGAGAGACC 597
Qy 2008 GTTCTCTGAGCGCTTACAGTGAACATCTTGGAGCGCGCCCTGTGCACCATCCAAACCTG 2067
Db 598 GTTCTCTGAGCGCTTACAGTGAACATCTTGGAGCGCGCCCTGTGCACCATCCAAACCTG 657

us-10-781-294-23.rng

Wed Feb 9 10:22:12 2005

Search completed: February 8, 2005, 16:38:28
Job time : 1538 secs

```
2068 ATAGAGTGTCTCTGTACCGAATGCCCTGGCAGCCGGGGGTGAAGCTGCTGTCTCAA 2127
      |||||||
658 ATAGAGTGTCTCTGTACCGAATGCCCTGGGAGCCGGGGGTGAAGCTGCTGTCTCAA 717
      |||||||
2128 GGACTCAGACACCCCAACTGCAAACTTCAGAACTGAGGCTGAAGAGTGCAGCATCTCC 2187
      |||||||
718 GGACTCAGACACCCCAACTGCAAACTTCAGAACTGAGGCTGAAGAGTGCAGCATCTCC 777
      |||||||
2188 AGCTCAGCCTGCGAGAGACCTCTCTGAGCTCTCATAGCCAAATAGAAATTTGACAAAGGATG 2247
      |||||||
778 AGCTCAGCCTGCGAGAGACCTCTCTGAGCTCTCATAGCCAAATAGAAATTTGACAAAGGATG 837
      |||||||
2248 GATCTCAGTGGCAACGGCGTTGATTCCCAAGGATGATGCTGTTTGCAGAGGCCCTGCGG 2307
      |||||||
838 GATCTCAGTGGCAACGGCGTTGATTCCCAAGGATGATGCTGTTTGCAGAGGCCCTGCGG 897
      |||||||
2308 CATCCCAAGTGCAGGCTGCAGATGATTCAGTTGAGGAAGTGTGAGTCCGGGGCT 2367
      |||||||
898 CATCCCAAGTGCAGGCTGCAGATGATTCAGTTGAGGAAGTGTGAGTCCGGGGCT 957
      |||||||
2368 TGTGAGGAGATGGCTTCTGTGCTGGGACCAACCCACATCTGTTGAGTTGAGCTTGAC 2427
      |||||||
958 TGTGAGGAGATGGCTTCTGTGCTGGGACCAACCCACATCTGTTGAGTTGAGCTTGAC 1017
      |||||||
2428 GGAATGCACTGGAGGATTTGGGCTGAGGTTACTATGCCAGGACTGAGGCAACCCAGTC 2487
      |||||||
1018 GGAATGCACTGGAGGATTTGGGCTGAGGTTACTATGCCAGGACTGAGGCAACCCAGTC 1077
      |||||||
2488 TGCAGACTAGGACTTTGTGCTGAGATCTGCCGCTCACTGCTGCTGCTGTGAGGAG 2547
      |||||||
1078 TGCAGACTAGGACTTTGTGCTGAGATCTGCCGCTCACTGCTGCTGCTGTGAGGAG 1137
      |||||||
2548 CTGGGCTCACTCTCAGTGTGAACCCAGAGCTGAGAGAGCTGAGACCTGAGCCTGAATGAG 2607
      |||||||
1138 CTGGGCTCACTCTCAGTGTGAACCCAGAGCTGAGAGAGCTGAGACCTGAGCCTGAATGAG 1197
      |||||||
2608 CTGGGGGACCTCGGGTGTGCTGTGTGAGGGCTCAGGCATCCCAAGTGCAGGCTC 2667
      |||||||
1198 CTGGGGGACCTCGGGTGTGCTGTGTGAGGGCTCAGGCATCCCAAGTGCAGGCTC 1257
      |||||||
2668 CAGACCTCGGGTGTGGGATCTGCCGCTGGGCTCTGCCGCTGTGAGGGTCTTTTCTGTG 2727
      |||||||
1258 CAGACCTCGGGTGTGGGATCTGCCGCTGGGCTCTGCCGCTGTGAGGGTCTTTTCTGTG 1317
      |||||||
2728 GTGCTCCAGGCCAACCAACCTCCGGGAGCTGGAATTTGAGTTTCAACGACCTGGGAGAC 2787
      |||||||
1318 GTGCTCCAGGCCAACCAACCTCCGGGAGCTGGAATTTGAGTTTCAACGACCTGGGAGAC 1377
      |||||||
2788 TGGGGCTGTGTTGTGCTGAGGGCTGCAACATCCGCTGAGAGCTCCAGAACTG 2847
      |||||||
1378 TGGGGCTGTGTTGTGCTGAGGGCTGCAACATCCGCTGAGAGCTCCAGAACTG 1437
      |||||||
2848 TGGCTGGATAGCTGTGGCTCAGCAAGGCTTGTGAAATCTTTTACTTCAACCTGGGG 2907
      |||||||
1438 TGGCTGGATAGCTGTGGCTCAGCAAGGCTTGTGAAATCTTTTACTTCAACCTGGGG 1497
      |||||||
2908 ATCAACAGACCTTGACCGACCTTTTACCTGACCAACGCGCTAGGGGACACAGTGTG 2967
      |||||||
1498 ATCAACAGACCTTGACCGACCTTTTACCTGACCAACGCGCTAGGGGACACAGTGTG 1557
      |||||||
2968 CGACTGCTTTGCAAGCGGCTGAGCCATCTTGGCTGCAAACTCCGAGTCTCTGGTTATTT 3027
      |||||||
1558 CGACTGCTTTGCAAGCGGCTGAGCCATCTTGGCTGCAAACTCCGAGTCTCTGGTTATTT 1617
      |||||||
3028 GGGATGGACCTGAAATAAAATGAACCCAGTAGTGGCAGCTTCGAGTAAACAACT 3087
      |||||||
1618 GGGATGGACCTGAAATAAAATGAACCCAGTAGTGGCAGCTTCGAGTAAACAACT 1677
      |||||||
3088 TATTTGGACATTTGGCTGCTGA 3108
      |||||||
1678 TATTTGGACATTTGGCTGCTGA 1698
      |||||||
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2005, 16:04:26 ; Search time 525 Seconds
(without alignments)
9686.758 Million cell updates/sec

Title: US-10-781-294-23
Perfect score: 3108
Sequence: 1 atgtacgaaccgagcag.....attggacattggtctgta 3108

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 398 | 12.8 | 4422 | US-09-388-221B-1 | Sequence 1, Appli |
| 2 | 398 | 12.8 | 4556 | US-09-388-221B-9 | Sequence 9, Appli |
| 3 | 395.8 | 12.7 | 4200 | US-09-388-221B-3 | Sequence 3, Appli |
| 4 | 395.8 | 12.7 | 4332 | US-09-388-221B-5 | Sequence 5, Appli |
| 5 | 395.8 | 12.7 | 4466 | US-09-388-221B-11 | Sequence 11, Appli |
| 6 | 299 | 9.6 | 1371 | US-08-910-731-1 | Sequence 1, Appli |
| 7 | 299 | 9.6 | 1371 | US-08-795-395-1 | Sequence 1, Appli |
| 8 | 293.2 | 9.4 | 1371 | US-08-910-731-7 | Sequence 7, Appli |
| 9 | 272.6 | 8.8 | 1386 | US-08-910-731-5 | Sequence 5, Appli |
| 10 | 256.6 | 8.3 | 1374 | US-08-910-731-3 | Sequence 3, Appli |
| 11 | 256.6 | 8.3 | 1374 | US-08-795-395-3 | Sequence 3, Appli |
| 12 | 85.8 | 2.8 | 257 | US-09-016-434-208 | Sequence 208, App |
| 13 | 83.2 | 2.7 | 376 | US-09-023-655-395 | Sequence 395, App |
| 14 | 77.4 | 2.5 | 678 | US-09-016-434-1011 | Sequence 1011, Ap |
| 15 | 68.2 | 2.2 | 4485 | US-09-949-016-3209 | Sequence 3209, Ap |
| 16 | 68.2 | 2.2 | 4485 | US-10-014-269-1 | Sequence 1, Appli |
| 17 | 68.2 | 2.2 | 4486 | US-10-014-269-33 | Sequence 33, Appli |
| 18 | 68.2 | 2.2 | 39936 | US-09-949-016-14951 | Sequence 14951, A |
| 19 | 61.2 | 2.0 | 1470 | US-09-099-041A-27 | Sequence 27, Appl |
| 20 | 61.2 | 2.0 | 1470 | US-09-245-281-27 | Sequence 27, Appl |
| 21 | 61.2 | 2.0 | 1470 | US-09-207-359B-27 | Sequence 27, Appl |
| 22 | 61.2 | 2.0 | 1470 | US-09-340-620A-27 | Sequence 27, Appl |
| 23 | 61.2 | 2.0 | 1470 | US-09-865-364-27 | Sequence 27, Appl |
| 24 | 61.2 | 2.0 | 2859 | US-09-099-041A-9 | Sequence 9, Appli |
| 25 | 61.2 | 2.0 | 2859 | US-09-245-281-9 | Sequence 9, Appli |
| 26 | 61.2 | 2.0 | 2859 | US-09-207-359B-9 | Sequence 9, Appli |
| 27 | 61.2 | 2.0 | 2859 | US-09-340-620A-9 | Sequence 9, Appli |

| | | | | | | |
|----|------|-----|-------|---|-------------------|-------------------|
| 28 | 61.2 | 2.0 | 2859 | 4 | US-09-865-364-9 | Sequence 9, Appli |
| 29 | 61.2 | 2.0 | 3080 | 3 | US-09-099-041A-25 | Sequence 25, Appl |
| 30 | 61.2 | 2.0 | 3080 | 3 | US-09-245-281-25 | Sequence 25, Appl |
| 31 | 61.2 | 2.0 | 3080 | 3 | US-09-207-359B-25 | Sequence 25, Appl |
| 32 | 61.2 | 2.0 | 3080 | 4 | US-09-340-620A-25 | Sequence 25, Appl |
| 33 | 61.2 | 2.0 | 3080 | 4 | US-09-865-364-25 | Sequence 25, Appl |
| 34 | 61.2 | 2.0 | 3382 | 3 | US-09-099-041A-7 | Sequence 7, Appli |
| 35 | 61.2 | 2.0 | 3382 | 3 | US-09-245-281-7 | Sequence 7, Appli |
| 36 | 61.2 | 2.0 | 3382 | 3 | US-09-207-359B-7 | Sequence 7, Appli |
| 37 | 61.2 | 2.0 | 3382 | 4 | US-09-340-620A-7 | Sequence 7, Appli |
| 38 | 61.2 | 2.0 | 3382 | 4 | US-09-865-364-7 | Sequence 7, Appli |
| 39 | 61.2 | 2.0 | 32042 | 3 | US-09-245-281-44 | Sequence 44, Appl |
| 40 | 61.2 | 2.0 | 32042 | 4 | US-09-340-620A-63 | Sequence 63, Appl |
| 41 | 58 | 1.9 | 2480 | 4 | US-09-064-199-15 | Sequence 15, Appl |
| 42 | 58 | 1.9 | 3393 | 1 | US-08-295-502-1 | Sequence 1, Appli |
| 43 | 58 | 1.9 | 3393 | 5 | PCT-US95-10691-1 | Sequence 1, Appli |
| 44 | 58 | 1.9 | 4346 | 4 | US-09-064-199-12 | Sequence 12, Appl |
| 45 | 58 | 1.9 | 4366 | 4 | US-09-064-199-14 | Sequence 14, Appl |

ALIGNMENTS

RESULT 1
US-09-388-221B-1
; Sequence 1, Application US/09388221B
; Patent No. 6818750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6818750el Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4422)
US-09-388-221B-1

| | | | | | |
|-----------------------|------|--|--------------------|------------|--------------|
| Query Match | | 12.8% | Score 398; | DB 4; | Length 4422; |
| Best Local Similarity | | 53.6% | Pred. No. 5.9e-96; | | |
| Matches 1082; | | Conservative 0; | Mismatches 855; | Indels 81; | Gaps 9; |
| Qy | 539 | CCGAGCCACCGCCACCGTGTCTATGCAAGCGCGGAGGATAGGCAAGTCCATGCTGG | 598 | | |
| Db | 971 | CCCAAGAACCTCGCATAGTCATCTGCAGGGGGCTGCTGGAATTGGGAAGTCAACACTGG | 1030 | | |
| Qy | 599 | CACACAGGTGATGCTGCTGCTGGCGGACCGGAGCTCTTCCAGGCGAGATTGATATC | 658 | | |
| Db | 1031 | CCAGGCAAGTGAAGGAAGCCCTGGGGGAGAGCCAGCTGTATGGGGACCGCTTCCAGCATG | 1090 | | |
| Qy | 659 | TCCTTACATCAACTGCAGCGAGATGAACAGAGTGCACGGAATGAGCATGCAAGACC | 718 | | |
| Db | 1091 | TCCTTACTTCAAGTGCAGAGAGCTGGCCAG---TCCAAGTGTGAGTCTCGCTGAGC | 1147 | | |
| Qy | 719 | TCATCTTCACTGCTGGCCTGAGCCAGCGCGCTCTCCAGAGCTCATCCAGTTC | 778 | | |
| Db | 1148 | TCATCGGAAAGATGGGACACCACTCCGGCTCCCATTAGACAGATCTGTCTAGGCCAG | 1207 | | |
| Qy | 779 | AGCGCTCTCTTTCATCATCGAGCTTCATGAGCTCAAGCTTCTTCCACGATCTC | 838 | | |
| Db | 1208 | AGCGCTGCTCTTTCATCTCGATGGTGTAGATGAGCCAGGATGGGTCTTTCAGGAGCCGA | 1267 | | |
| Qy | 839 | AGGACCTGCTGCTGCTGCTGGGAGAGAAACGGCCACCGAGCTCTCTTCTTAACAGCT | 898 | | |
| Db | 1268 | GTTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGCCGGGGATGCACTGCTGGGCGATT | 1327 | | |
| Qy | 899 | TAATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGCGCCACGGCTT | 958 | | |

1328 TGC TGGGAAAAC TATACTTCCGAGGCATCTTCTCTGATCAGCGCTCGGACCACAGCTC 1387

959 TGGAGAGCTCACCGTCTGCTGGAGCACCCAGGCATGTGGAGATCCTGGGCTTCTCTG 1018

1388 TGCAGAACCTCAATTCCTTCTTGGAGCAGGCACGTTGGGTAGAGTCTCTGGGTTCTCTG 1447

1019 AGSCAGAAAAGGAAGAACTTCTTCAACAGTATTTCCAAATPGCAGACGCGGGGCAAG 1078

1448 AGTCAGCAGGAAGGAATATTTCTACAGATATTTACAGATGAAGGCAAGCAATTAGAG 1507

1079 TCTTCAATTAAGTGGAGGACAAGAGCTCTCTTCCATCATGTGCTTGGTCCCTGGTGT 1138

1508 CCTTTAGTTGGTCAAAATCAAAACAAGAGCTCTGGGCTCTGTGCTTGTGCCCTGGGTGT 1567

1139 GCTGGGTGTGTATCACTGCCCTCCAGCAGCAGCTGGAGGTGGGGGCTGTGAGACAGA 1198

1568 CTTGGCTGGCCGTGCACTTGCCTGATGCAGCAGATGAAGCGAAGMAAACTCACACTGA 1627

1199 CGTCCAGGACCAACTGTCAGTGTATCATGCTCTACTCTGAGTCTGATGCAACCCCAAGC 1258

1628 CTTTCCAAGACCAACAACCCCTCTGTCTACATTTACTTTCGCCAGGCTCTCCAGCTCAGC 1687

1259 CGGGGCGCCGCGCCTCCAGCCCCACCCAAACAGAGAGGGTTGTCTCTTTGGCGGAG 1318

1688 CATTGG-----GACCCAGCTCAGAGACCTCTGCTCTCTGCTGCTG 1729

1319 ATGGGCTCTCGAATCAGAAAATCTATTTGAGAGCAGGACCTCGGAGCAGCGCCTAG 1378

1730 AGGGCATCTGGCAAAAAGAACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTAG 1789

1379 ACGGGGAAGACGTCTCTGCCCTCCCTCAAATGAACATCTTCAGAAAGACATCAACTGTG 1438

1790 ATGGGGCCATCACTCCACCTTCTTGAAGATGGGTATCTTCAAGACACCCCATCCCTC 1849

1439 AGAGGTACTACAGTTCATCCATTCAGTGTTCAGGAATTCCTTTGAGCTATGACTATA 1498

1850 TGAG---CTACAGCTTCATTCACTCTGTTTCCAGAGTTCCTTTCGACGAAATGTCTCATG 1906

1499 TCCTGGACGAGGGGAGCGGGGCGGACAGCAGGACGTGACAGGCTGTGTGACCG 1558

1907 TCTTGGAGGATGAAGGGAGAGGTAAACATTTCAATTCATATAGATTTGGAAAAAGA 1966

1559 AGTACGCGTTTTCTGAAAGGAGCTTCCT-----GGCACTCACAGCGCTTCCTGTTTG 1612

1967 CGCTAGAGCATATGGAATACATGGCCTGTTTGGGGCATCAACCAACGTTTCTTATTTGG 2026

1613 GACTCTGAAACAGAGGACACGAGGCCACTCTGGAGAGAGTCTCTCTCGAAAGCTCTGC 1672

2027 GCCTGTTAAGTGATGAGGGGAGAGAGATGGGAACATCTTTCACTGCCGCTGCTC 2086

1673 CGCACATCAAGATGACCTGTTGCAGTGGATCCAAAGCAAAAGCTCAGAGCGACGGCTCCA 1732

2087 AGGG-----GAGGAACCTGATGCAGTGGGTCCGCTCCCTGCAGCTGCT----- 2129

1733 CCCTCGAGAGGGCTCCTTGGAGTTCACGTCTGTGACAGATCCAGGAGGAGGAGT 1792

2130 -GCTCGAGCCACACTCTCTGAGTCCCTCCACTGCTTGTACAGACATCTGGAAACAACGT 2188

1793 TTATCCAGCAGGCCCTGAGCCACTTCCAGGTGATCTGGTCAAGCAATTCCTCCAAGA 1852

2189 TCCTGACACAGTGATGCCCATTTCCGAAGAAA--TGGGCATGTGTGTAGAAACAGACA 2245

1853 TGGAGCACAATGCTCTCTGTTCTGTCTGAAGCGCTGCAGAGCGCCAGGTCTGCACCT 1912

2246 TGGAGCTCTTAGTGTGCATCTTCTGCATTTAAATTAAGCCGCCACGTGAAGAGCTTCAGC 2305

1913 TGTATGGGCCCACTACAGCGCGGAGGAGACCGCGGAGGTGCTCCGAGGAGCGC 1972

2306 TGATT-----GAGGGCAGGACACAGATCAACATGATGAGCC 2341

1973 ACACGCTGTTGTGACGTCTAGACAGAGGACCGTTCTCTCTGGAGCGCTACAGTGAAC 2032

| | | | |
|--|------|---|------|
| Db | 2342 | CCACCATGGTAGTCCCTGTTTCAG-----GTGGGTCCCAAGTCACAGATGCCTATTGCGCAGA | 2395 |
| Qy | 2033 | ATCTGGCAGGGCCCTGTGCAACCAATCCAAACCTGATAGAGCTGTCTCTGTACCGAAATG | 2092 |
| Db | 2396 | TTCTCTTTCCTCGTCTCAAGGTCACAGAAACCTGAAGAGCTGGACCTAAGTGGAAACT | 2455 |
| Qy | 2093 | CCCTGGGCGACCCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCAACTGCAAC | 2152 |
| Db | 2456 | CGCTGAGCCACTCTGCAGTGAAGAGTCTTTGTAAGAACCCCTGAGACGCCCTCGCTGCCTCC | 2515 |
| Qy | 2153 | TTCAGAACTGAGGCTGAGAGGTGCCGCATCTCCAGCTCAGCTGCGAGGACCTCTCTGTG | 2212 |
| Db | 2516 | TGGAGACCTTCGCGTTGGCTGTGGCTTACAGCTGAGACTGCAAGGACCTTGCT | 2575 |
| Qy | 2213 | CAGCTCTCATAGCCAATAAGAAATTTGACAAGGATGGATCTCAGTGGCAACGGCGTTGGAT | 2272 |
| Db | 2576 | TTGGGCTGAGGCCAACACGAGACCCCTGACGAGCTGGACCTGAGCTTCAATGTGCTCACGG | 2635 |
| Qy | 2273 | TCCCAGGCAATGATGCTGCTTTGGAGGGCTGCGGCATCCCAAGTGCAGGCTGCAGATGA | 2332 |
| Db | 2636 | ATGCTGGAGCCAAACACCTTTTGCAGAGACTGAGACAGCGGAGCTCAAGCTACAGCGAC | 2695 |
| Qy | 2333 | TTCAAGTTGAGGAAGTGTCAGCTGGAGTCCGGGGCTTTGTTCAGGAGATGGCTTCTGTGCTCG | 2392 |
| Db | 2696 | TGCAGCTGGTCAGCTGTGGCTCACGCTCTGACTGCTGCCAGGACCTGGGCTCTGTGCTTGA | 2755 |
| Qy | 2393 | GCACCAACCCACATCTCTGGTTGAGTTGGACCTGACGGAATGCACTGGAGGATTTGGGCC | 2452 |
| Db | 2756 | GTGCCAGCCCCAGGCTCTGAAGGAGCTAGACCTGACGAGCAAAACCTGGATGACGCTGGCG | 2815 |
| Qy | 2453 | TGAGGTTACTATGCCAGGACTGAGGCACCCAGTCTTCAGACTACGGACTTTTGTGGCTGA | 2512 |
| Db | 2816 | TGGGACTGCTGTGAGGGGCTCAGGCATCTTGCTGCAAACTCATACGCTGGGGCTGG | 2875 |
| Qy | 2513 | AGATCTGCCGCTCACCTGCTGCTGCCCTGTGACGAGCTG | 2550 |
| Db | 2876 | ACCAGACAACTCTGAGTGATGAGATGAGCGGAACTG | 2913 |
| RESULT 2 | | | |
| US-09-388-221B-9 | | | |
| ; Sequence 9, Application US/09388221B | | | |
| ; Patent No. 6818750 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Reed, John C. | | | |
| ; TITLE OF INVENTION: No. 6818750el Card Proteins Involved in Cell Death Regul | | | |
| ; FILE REFERENCE: P-LJ 3650 | | | |
| ; CURRENT APPLICATION NUMBER: US/09/388,221B | | | |
| ; CURRENT FILING DATE: 1999-09-01 | | | |
| ; NUMBER OF SEQ ID NOS: 30 | | | |
| ; SOFTWARE: Patentin Ver. 2.0 | | | |
| ; SEQ ID NO 9 | | | |
| ; LENGTH: 4556 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Artificial Sequence | | | |
| ; FEATURE: | | | |
| ; NAME/KEY: CDS | | | |
| ; LOCATION: (1)..(4365) | | | |
| ; FEATURE: | | | |
| ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic | | | |
| ; OTHER INFORMATION: Construct | | | |
| US-09-388-221B-9 | | | |
| Query Match | | | |
| Best Local Similarity 12.8%; Score 398; DB 4; Length 4556; | | | |
| Matches 1082; Conservative 53.6%; Pred. No. 6e-96; | | | |
| 9 | | | |
| Gaps 81; | | | |
| Qy | 539 | CCGAGCACCGCGCACCGCTGGTTCATGCAAGGGCGCGCAGGGATAGCGAAGTCCATGCTGG | 598 |
| Db | 971 | CCCAAGAACTTCGCATGATCATCTGACGGGGCTGCTGGAATTGGGAAGTCAACACTGG | 103 |
| Qy | 599 | CACACAAGGTGATGCTGGACTGGCGGACGGGAGCTCTTCCAAAGCGACATTTGATTATC | 658 |

Db 1031 CCAGCAGGTGAAGAACCTCGGGGAGAGGCCAGCTGTATGGGGAACGCTTCAGCATG 1090
Qy 659 TCCTTCAATCAATCACTGCAGGAGATGAACAGAGCTGCCACCGAATGACAGCATGCAAGACC 718
Db 1091 TCCTTCACTTCACTGCAGAGAGCTGGCCAG---TCCAAGTGGTGAATCTCGCTGAGC 1147
Qy 719 TCATCTTCAGCTGCTGGCTGAGCCGACGCGCTCTCCAGGAGCTCATCCGAGTTCGCG 778
Db 1148 TCATCGGAAAAGATGGGACACCACTCGGCTCCCATTTAGACAGATCTGTCTTAGGCCAG 1207
Qy 779 AGCGCTCTTTTCATCATCAACGCTTCGATGAGCTCAAGCTTCTTTCCACATCTTC 838
Db 1208 AGCGCTGCTCTTCATCTCCTGATGTGATGAGCCAGGATGGGTCTTGCAGAGCCGA 1267
Qy 839 AGGACCTCTGTGCTCTGCTGGGAGGAGAAACGCGCCACGAGCTGCTTCTTAACAGCT 898
Db 1268 GTTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGAGCGCGGATGCACTGCTGGGAGTT 1327
Qy 899 TAAATCGGAAGAAGCTGCTCCTGAGCTATCTTTGCTCATCACCAACGCGCCACGCTT 958
Db 1328 TGCTGGGGAAAACATACTTCCGAGGCACTCTTCTGATCAACGCTCGGACACAGCTC 1387
Qy 959 TGGAGAGCTCCACGCTGCTGTCGAGGACCCAGGCACTGTGGAGATCTGGGCTTCTCTG 1018
Db 1388 TGCAGAACCTCATCTCTTCTTTGGAGCAGGACGCTGGGTGAGAGGTCTCGGGTCTCTG 1447
Qy 1019 AGGCAGAAAGGAAGAAATCTTCTACAAGTATTTCCAAATGTCAGAGCAGCGCGGCCAAG 1078
Db 1448 AGTCAGCAGAGAGGAATATTTCTACAGATATTTCAAGATGAAGGACGAATTAGAG 1507
Qy 1079 TCCTTCAATACGTGAGGGAACAAGCTCTCTTCAACATGTGCTTCGTCCTCCCTGCTGT 1138
Db 1508 CTTTGTAGTGGTCAATCAAAACAAGAGCTCTGGGCCCTGTCTTGTGCTCGCTGGTGT 1567
Qy 1139 GCTGGGTGGTGTACCTGCTCAGCAGAGCTGGAGGTGGGGGCTGTGTAGACAGA 1198
Db 1568 CTTGGCTGGCTGCACTTGGCTGTATGTCAGCAGATGAAGCGGAAGAAACTCACTGA 1627
Qy 1199 CGTCAGGACCAACACTGCACTGTACATGCTCTACCTGCTGAGTCTGATGCAACCCCAAGC 1258
Db 1628 CTTTCAAGACCAACAACCTCTGTCTACATTAACCTTGGCCAGGCTCTCAGCTCAGC 1687
Qy 1259 CGGGGGCCCCGCTCTCAGCCCCCACCACCAAGAGAGGTTGTCTCTTGGCGGAG 1318
Db 1688 CATTTGG-----GACCCAGCTCAGAGACCTCTGCTCTCTGCTGCTG 1729
Qy 1319 ATGGGCTCTGAAATCAGAAATCTATTTAGAGAGCAGGACCTCGGAAGACGCGCTAG 1378
Db 1730 AGGGCATCTGGCAAAAAGAACCTTTTCACTCCAGATGACCTCAGGAAGCATGGGTAG 1789
Qy 1379 ACGGGGAAGCTCTCTGCTTCTCAACATGAACATCTCCAGAGGACATCAACTGTG 1438
Db 1790 ATGGGCAATCATCTCAACCTCTTGAAGATGGGTATTTCTTCAAGAGACCCCATCCCTC 1849
Qy 1439 AGAGGTACTACAGCTTCATCACTTGAGTTTCCAGGAATCTTTGACGCTATGTACTATA 1498
Db 1850 TGAG--CTACAGCTTCATTCACTCTGTCTTCCAGAGATCTTTGACAGAAATGTCCTATG 1906
Qy 1499 TCCTGGACGGGGAGGGCGGGGAGGCCCCAGACGAGAGCTGACACAGGCTGTGTGACG 1558
Db 1907 TCTTGGAGGATGAGAAGGGGAGAGGTAAACATTTCTAATTGATCATAGATTTGGAAGA 1966
Qy 1559 AGTACGCTTTTCTGAAGGAGCTTCCT-----GGCACTCACGAGCGCTTCCTGTTTG 1612
Db 1967 CGCTAGAAGCATATGGAATACATGGCTCTGTTGGGGCATCAACACAGCTTCTCTATTTG 2026
Qy 1613 GACTCTCTGAACAGGAGACCAAGGACCACTGGAGAGAGTCTCTGCTGGAAGTCTCGC 1672
Db 2027 GCCTGTTAAGTGATGAGGGGAGAGAGAGATGAGAGACATCTTTCACTGCGCGCTGCTC 2086
Qy 1673 GGCACATCAAGATGACCTGTGATGGATCCAAAGCAAAAGCTCAGAGGACGGCTCCA 1732
Db 2087 AGGG-----GAGGAACCTGATGAGTGGGTCCGCTCCCTGCACTGCT----- 2129

Qy 1733 CCCTGCAGCAGGGCTCCTTGGAGTTCTTACGCTGCTTTGTACGATCCAGGAGGAGGT 1792
Db 2130 -GCTGCAGCCACACTCTCTGAGTCCCTCCTCCTGCTTGTACGAGACTCGGAACAAACGT 2188
Qy 1793 TTATCCAGCAGGGCTGAGCCATTTCCAGGTGATGTTGGTTCAGCAACATTTGCTCTCAAGA 1852
Db 2189 TCCTGACACAAGTGATGGCCCATTTTGAAGAAA---TGGGATGTGTGTGAACACAGACA 2245
Qy 1853 TGAGACATATGTTCTCTCTGCTGAAGCGCTGCAGAGCGCCAGTGTCTGCACT 1912
Db 2246 TGGAGCTCTTAGTGTGCACTTTCTGCATTAATTCAGCCCGCATGTAAGAGCTTCAGC 2305
Qy 1913 TGTATGCGCCACCTACAGCGCGGACCGGGGAAGACCGCGGAGGTGCTCCGCGAGAGCGC 1972
Db 2306 TGATT-----GAGGGCAGGACGACAGATCAACATGGAGCC 2341
Qy 1973 ACACGCTTTGGTGCAGCTCAGACGAGAGAGCGTTCCTGTGTGAGCGCTTACAGTGAAC 2032
Db 2342 CCACCATGTGTGCTCTGTTTCCAG-----GTGGGTCCAGTTCACAGATGCTTATTTGGCAGA 2395
Qy 2033 ATCTGGCAGCGGCTGTGCACCAATCCAAACCTGATAGAGCTGCTCTGTACCGAAATG 2092
Db 2396 TTCTTCTCCGCTCTCAAGGTCAACAGAAACCTGAAGAGCTGGACCTAAGTGGAACT 2455
Qy 2093 CCCTGGCAGCCGCGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCAACTGCAAAAC 2152
Db 2456 CGCTGAGCCACTCTGCAGTGAAGAGTCTTTGTAAAGCCCTGAGACGCGCTCGCTGCCTCC 2515
Qy 2153 TTCAAGAACCTGAGGCTGAAGAGGTGCCGATCTCTCAGCTCAGCTCAGGAGACCTCTCTG 2212
Db 2516 TGGAGACCTTGCCTGGCTGTGGCTTGCCTTCAACAGCTGAGGACTGCAAGGACCTTGCCT 2575
Qy 2213 CAGCTCTATAGCCAAATAAGAAATTTGAACAGGATGATCTCAGTGGCAACGCGCTTGGAT 2272
Db 2576 TTGGGCTGAGAGCCAAACAGACCTGACCGAGCTGGACCTTGAGCTTCAATGTGCTCAGG 2635
Qy 2273 TCCAGGACATGATGCTGCTTTTGGAGGCGCTGGGCACTCCCGAGCTGAGGCTGAGATGA 2332
Db 2636 ATGCTGGAGCCAAACACTTTTGCAGAGACTGAGACAGCGCGCTGCAAGCTACAGCGAC 2695
Qy 2333 TTCAAGTTGAGGAAGTGTGAGTGGAGTCCGGGGCTTGTTCAGGAGATGGCTTCTGTGCTG 2392
Db 2696 TGCAGCTGTGTGAGTGTGGCTTCACTGCTGACTGCTGCCAGGACCTGGCTCTGTGCTTA 2755
Qy 2393 GCACCAACCACTCTGTTGAGTTGACCTGACAGGAAATGCACTGAGGATTTGGGCC 2452
Db 2756 GTGCCAGCCCGCTGAAGAGGCTAGCTGCTGAGCAGAAACACCTGGATGACGTTGGCG 2815
Qy 2453 TGAGGTTACTATGCAAGGACTGAGGCAACCCAGTCTGCAGACTACGGAATTTTGGGCTGA 2512
Db 2816 TGGCACTGCTCTGTGAGGGCTCAGGATCTCTGCCTGCAAACTCATACGCTGGGGCTGG 2875
Qy 2513 AGATCTGCCCGCTCAGCTGCTGCTGCTGTGACGAGCTG 2550
Db 2876 ACCAGAACACTCTGAGTGTATGATGAGGAGGAACTG 2913

RESULT 3

US-09-388-221B-3
; Sequence 3, Application US/09388221B
; Patent No. 6818750
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. 6818750el Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 4200
; TYPE: DNA

Wed Feb 9 10:22:12 2005

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4197)
; US-09-388-221B-3

Query Match      12.78; Score 395.8; DB 4; Length 4200;
Best Local Similarity 54.08; Pred. No. 2.2e-95;
Matches 1060; Conservative 0; Mismatches 822; Indels 81; Gaps 9;

QY 539 CCGAGCCACCGCGCACCGTGGTTCATGCAAGCGCGGAGGATAGCAAGTCCATGCTGG 598
DB 971 CCCAAGAACCTCGCATAGTCATAGTCAGGGGGCTGTGGAATTGGGAAGTCAACACTGG 1030
QY 599 CACACAGGTGATCTGGACTGGCGGACCGGAGGCTCTTCCAAGCAGATTGATTATC 658
DB 1031 CCAGCGAGGTGAAGAAAGCCCTGGGGGAGAGCCAGCTGTATGGGACCGCTTCCAGCATG 1090
QY 659 TCTTCTACATCAACTGCGAGGAGATGAACAGAGTGCACCGAATGCGAGATGCAAGACC 718
DB 1091 TCTTCTACTTCAGCTGCGAGAGCTGGCCAG---TCCAAGGTGGTGAATCTCGCTGAGC 1147
QY 719 TCATCTTCAGTGTGGCTGAGCCAGCGGCTCTCCAGGAGCTCATCCGAGTTCCTCG 778
DB 1148 TCATCGGAAGAGATGGGACACCCACTCGGCTCCCATATGACAGATCTCTGTAGGCCAG 1207
QY 779 AGCGCTCTCTTTCATCATGACGCTTCGATGAGCTCAAGCCCTTCTTCCAGATCCTC 838
DB 1208 AGCGCTGCTCTTTCATCTCGATGGTGTAGATGAGCCAGGATGGTCTTGCAGAGCCGA 1267
QY 839 AGGACCCCTGFGCTCTGCTGGGAGAGAAACGGGCCACCGGAGCTCTTCTTAAACGCT 898
DB 1268 GTTCTGAGCTCTGCTGCACTGGAGCCACGACAGCGCGGCGATGCACTGTGGGCGATT 1327
QY 899 TAACTCGGAGAGCTGCTCCTGAGCTATCTTGTCTCATCACACAGCGGCCACGCTT 958
DB 1328 TGTGGGGAAGAACTATCTTCCGAGGATCTTCTCTGATCACGCTCGGACACAGCTC 1387
QY 959 TGGAGAAGCTCCACCGTCTGCTGGAGCACCCAGGCGATGTGGAGATCTTGGGCTTCTG 1018
DB 1388 TGCAGAACCTCATCTCTTCTTGGAGCAGGACGTTGGGTAGAGTCTCGGGTTCCTG 1447
QY 1019 AGCGAAGAGGAGGATCTTCAAGTATTTCCAAATGCAAGCAGCGCGGCCCAAG 1078
DB 1448 AGTCCAGCAGGAAGGAATTTCTACAGATATTTACAGATGAAGGCAAGCAATTAGAG 1507
QY 1079 TCTTCAATTACGTGAGGGAACAAGGCTCTCTTCAACATGCTGTGTGTCCTCGGTGT 1138
DB 1508 CCTTAGTTGGTCAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCTGGGTGT 1567
QY 1139 GCTGGGTGGTGTATCTGCTCCAGCAGCTGGAGGTGGGGGCTGTTGACACAGA 1198
DB 1568 CTTGGCTGGCTGCACTTGCCTGATGACAGATGAAGCGGAAGGAAATCTACACTGA 1627
QY 1199 GTCCAGGACACCACTGCACTGTACATGCTCTTACCTGTGAGTCTGATGCAACCAAGC 1258
DB 1628 CTTCCAGACACCAACCACTCTGTCTACATTAACCTTGCCTCCAGGCTCTTCCAGCTCAG 1687
QY 1259 CGGGGGCCCGCGCTCCAGCCGCCACCCAAACAGAGAGGTGTGCTCTTGGGGGAG 1318
DB 1688 CATTTGG-----GACCCAGCTCAGAGACCTCTCTCTTCTGGCTGTG 1729
QY 1319 ATGGGCTTGAATCAGAAATCTATTGAGGAGCAGGACCTCCGGAAGCAGCGGCTAG 1378
DB 1730 AGGGCATCTGGCAAAAAAAGACCTTTTCAGTCCAGATGACCTCAGGAGCATGGGTAG 1789
QY 1379 ACGGGAAGAGCTCTCTGCTCTTCTCAACATGAACATCTTCCAGAGGACATCACTGTG 1438
DB 1790 ATGGGGCCATCATCTCCACCTCTTGAAGATGGGTATTTCTCAAGAGCACCCCATCCCTC 1849
QY 1439 AGAGTACTACAGCTTCTACCTTGTGTTTCCAGGAATCTTTGACGCTATGTACTATA 1498
DB 1850 TGAG---CTACAGCTTCATTACCTCTCTGTTTCCAGAGTCTTTTGACAGCAATGCTCTATG 1906

```

RESULT 4

US-09-388-221B-5
; Sequence 5, Application US/09388221B

```

QY 1499 TCCTGACAGGGGGAGGGCGGCGGAGGCCAGAGCCAGGACCTGACAGGCTGTTGACCG 1558
DB 1907 TCTTGGAGGATGAGAAGGGGAGAGGTAAACAATTCATTTGCATCATAGATTTGGAAGA 1966
QY 1559 AGTACGGCTTTTCTGAAAGGAGCTTCTCT-----GGCACTACACGCGCTCTCTGTTG 1612
DB 1967 CGCTAGAGCATATGGAATACATAGGCTGTGTTGGGCACTCAACACACGCTTTCTTATTGG 2026
QY 1613 GACTCTGAAACGAGGAGACCAAGAGGACCACTGGAGAGAGTCTCTGCTGGAAGGTCTCGC 1672
DB 2027 GCCTGTTAAGTGTAGTGGGGGAGAGAGATGAGAAACATCTTTCACTGCGCGCTGCTC 2086
QY 1673 CGCACATCAAGATGAGCCTGTTCAGTGTGATCCAAAGCAAGCTCAGAGCGAGCGGTCCA 1732
DB 2087 AGGG-----GAGGAACCTGATGCAAGTGGTCCCGTCCCTGCACTGC----- 2128
QY 1733 CCCTGACAGCGGCTCTTGGAGTTCCTGAGCTCTTGTACGCTGCTTGTACGAGATCCAGGAGGAGT 1792
DB 2129 TGGTGCAGCACACTCTCTGAGTCCCTCCACTGCTTGTACGAGACTCGGAACAAACGCT 2188
QY 1793 TTATCCAGCGGCTCTGAGCCACTTCCAGTGTGATCGTGTGTCAGCAACATTTGCCCTCCAAGA 1852
DB 2189 TCCTGACACAAAGTANTGGCCCATTTTCGAAGAA---TGGCATGTGTGTAAGAAACAGACA 2245
QY 1853 TGGAGCACATGGTCTCTCTGTTCTGTGAAGCGCTCCAGAGCGGCCAGGTCTGCACT 1912
DB 2246 TGGAGCTCTTAGTGTGCACTTTCTGCAATTAATTCAGCCGCGCTGGAAGCTTTCAGC 2305
QY 1913 TGTATGGCGCCACCTACAGCGCGGAGAACCGCGAGGTGTCTCCGAGGAGCGC 1972
DB 2306 TGATT-----GAGGGCAGGACGACAGATCAACATGGAGCC 2341
QY 1973 ACACGCTGTGTGTCAGCTCAGACCAAGAGGACCGTTCTGCTGGAAGCTTACAGTGAAC 2032
DB 2342 CACACATGTTAGTCTCTGTTTTCAG-----GTGGGTCCCAGTCAAGATGCTTATTTGGCAGA 2395
QY 2033 ATCTGCGAGCGGCTCTGTGTCACCAATCCAAACCTGATAGCTGTCTGTACCGAATG 2092
DB 2396 TTCTCTTCTCTCTCTCAAGGTCACCAAGAACTCTGAGAGCTTGGACCTAAGTGGAACT 2455
QY 2093 CCCTGGGACGCGGGGGTGAAGCTGTCTGTCAAAGGACTCAGACACCCCAACCTGCAAC 2152
DB 2456 CGCTGAGCACACTCTGCACTGAGAGTCTTTGTGAAGACCTGAGACGCTCGCTGCTCC 2515
QY 2153 TTCAAGAACCTGAGGCTGGAAGAGTGGCGCATCTCCAGCTCAGCTGCGAGGACCTCTCTG 2212
DB 2516 TGGAGACCTTGGCGTGGCTGTGGCTCACAGCTGAGGACTGCAAGGACCTTGCT 2575
QY 2213 CAGCTCTCATAGCCAAATAAGAAATTTGACAAAGGATGGATCTCAGTGGCAACGCGCTTGGAT 2272
DB 2576 TTGGGCTGAGAGCCAAACCAAGCCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCACGG 2635
QY 2273 TCCAGAGCATGATGCTGCTTTTGGAGGGCTGTGGGATCCCAAGTGCAGGCTGCAAGTGA 2332
DB 2636 ATGCTGGAGCCAAACACCTTTTGGCAGAGACTGAGACAGCGAGCTGCAAGCTACAGCGAC 2695
QY 2333 TTCAATTGAGAGTGTACGCTGAGTCCGGGCTTGTTCAGGAGATGGCTTCTGTGCTCG 2392
DB 2696 TGCAGCTGGTCACTGTGGCTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2755
QY 2393 GCACCAACCAACATCTGTTGAGTGGACCTGGAACAGAAATGCACTGAGGAGATTTGGGCC 2452
DB 2756 GTGCGAGCCCGAGCTTGAAGGAGCTAGACCTGACGAGCAACAACTGGATGACGTTGGCG 2815
QY 2453 TGAGGTACTATGTCAGGAGTGGGACCGAGCTCTGCACT 2495
DB 2816 TGCAGTGTCTGTGTGAGGGGCTCAGGCACTCTGCTGCTGCAAACT 2858

```

```

; Patent NO. 6818750
;
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: NO. 6818750a1 Card Proteins Involved in Cell Death Regulation
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221B
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (4332)
; US-09-388-221B-5

```

Query Match 12.7%; Score 395.8; DB 4; Length 4332;
Best Local Similarity 54.0%; Pred. No. 2.3e-95;
Matches 1060; Conservative 0; Mismatches 822; Indels 81; Gaps 9;

| | | |
|------|---|------|
| 539 | CCGAGCCACCGCGCACCGTGGTCTATGAAAGCGCGGAGGATAGGCAAGTCCATGCTGG | 598 |
| Db | | |
| | | |
| 971 | CCCAAGAACCTTCGATAGTATCATCTGCAGGGGGTCTGGAAATGGGAGATCAACACTGG | 1030 |
| Db | | |
| | | |
| 599 | CACACAGGTGATGCTCGACTGGCGGACGGAGACTCTTCCAAGGCAGATTGTATTATC | 658 |
| Db | | |
| | | |
| 1031 | CCAGGCAGGTGAAGAAAGCCCTGGGGGAGAGGCCAGCTGTATGGGGACCGCTTCCAGCATG | 1090 |
| Db | | |
| | | |
| 659 | TCCTTCTACATCAACTGCAGGGAGATGAACACAGAGTGCACGGAAATGCAGCATGCAAGACC | 718 |
| Db | | |
| | | |
| 1091 | TCCTTCTACTTCAGCTGCAGAGAGCTGGCCCAAG---TCCAAGTGGTGAGTCTCGCTGAGC | 1147 |
| Db | | |
| | | |
| 719 | TCATCTTTCAGTGTGTGCGCTGAGCCCAAGCGCGCTCTCCAGAGAGCTCATCCGAGTTCCCG | 778 |
| Db | | |
| | | |
| 1148 | TCATCGAAAGATGGACACCACTCCGGCTCCCATTTAGACAGATCCTGTCTAGGCCAG | 1207 |
| Db | | |
| | | |
| 779 | AGGCGCTCTTTTCATCATCGAGGGCTTCAGATGAGCTCAAGCCCTCTTTTCCACAGATCCTC | 838 |
| Db | | |
| | | |
| 1208 | AGCGCTGTCTTTCATCCTCGATGGTGTAGATGAGCCAGAGATGGGTCTTTCGAGAGCCGA | 1267 |
| Db | | |
| | | |
| 839 | AGGGACCTCGTGTCTGCTGGGAGGAGAAACGGCCACCGAGAGCTGCTTCTTTAAACAGCT | 898 |
| Db | | |
| | | |
| 1268 | GTTCTGAGCTCTGTCTGCACTGGAGCCAGCCACAGCCGGCGGATGCACTGCTGGCGAGTT | 1327 |
| Db | | |
| | | |
| 899 | TAATTCGGAAGAGCTGCTCCTGAGCTATCTTTGTCTATCAACAACGGGCCACGGGCTT | 958 |
| Db | | |
| | | |
| 1328 | TGCTGGGGAAAACTATATCTTCCCAGGACATCTCTTCTGTATCAGCGCTCGGACCAACAGCTC | 1387 |
| Db | | |
| | | |
| 959 | TGAGAGAGCTCCACCGTCTGCTGGAGACACCCACAGGCATGTGGAGATCCTGGGCTTCTCTG | 1018 |
| Db | | |
| | | |
| 1388 | TGCAGAACCTCATCTCTTTTGAGCAGGCACGTTGGGTAGAGGTCTCGGGTCTCTCTG | 1447 |
| Db | | |
| | | |
| 1019 | AGGCAGAAAGGAAGGAATATCTTCTCAAGTATTTTCCAAATGACAGAGCAGCGCGGCCAAG | 1078 |
| Db | | |
| | | |
| 1448 | AGTCCACGAGGAAGGAATATTTCTACAGATATTTTCACAGATGAAGACCAAGCAATTAGAG | 1507 |
| Db | | |
| | | |
| 1079 | TCCTTAATTAAGTGAGGGAACAACGAGCCTCTCTTTCACAATGTGCTTCGTCCCTCGTGT | 1138 |
| Db | | |
| | | |
| 1508 | CCTTTAGTTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTTGTGCGCTGGGTGT | 1567 |
| Db | | |
| | | |
| 1139 | GCTGGGTGGTGTGTAACCTGCTCCAGCAGCAGCTGGAGGTTGGGGGCTCTGTAGACAGA | 1198 |
| Db | | |
| | | |
| 1568 | CCTGGCTGGCCGTGCACTTGGCCTGTATGAGCAGATGAAGCGGAAGGAAAAACTCACTGA | 1627 |
| Db | | |
| | | |
| 1199 | CGTCCAGGACCAACCACTGCAGTGTAATGCTTACCTCTGAGTCTGATGCAAGCCCAAGC | 1258 |
| Db | | |
| | | |
| 1628 | CTTCCAAGACCAACCAACACCTCTGCTCTACATTAACCTTGCCCGGCTCTCCAAGTCAAGC | 1687 |
| Db | | |
| | | |
| 1259 | CGGGGGCCCCGGCCTCCAGCCCCCAACCAACAGAGAGGGTTGTCTCTCTTGGCGGAG | 1318 |
| Db | | |
| | | |
| 1688 | CATTGG-----GACCCAGCTCAGACACTCTGCTCTCTCTGTGCTGCTGCTG | 1729 |

| | | | |
|------|----|---|------|
| 1319 | Qy | ATGGGCTCTGGAAATCAGAAAAATCCTATTATTAGAGAGCAGGACCTCCGGAAGACACGGCCTAG | 1378 |
| 1730 | Db | AGGGCATCTGGCAAAAAAAGACCCCTTTTCAGTCCAGATGACCTTCAGGAAGCATGGGGTTAG | 1789 |
| 1379 | Qy | ACGGGAAGACGCTCTGCTGCTTCCCTCAACATGAACATCTTCCAGAAGGACATCAACTGTG | 1438 |
| 1790 | Db | ATGGGGCCATCATCTCCACCTTCTTGAAATGGGTATCTTCAAGAGCACCCCATCCCTC | 1849 |
| 1439 | Qy | AGAGGTACTACAGCTTCATCCACTTCAGTTTTCAGGAAATCTTTTGACGTATCTACTATA | 1498 |
| 1850 | Db | TGAG--CTTACAGCTTCATTCACCTCTGTTTCCAGAGTCTTTTGACGAAATGCTCTATG | 1906 |
| 1499 | Qy | TCTTGACAGAGGGGAGGGCGGGCAGGCCAGACAGAGACGTGAACAGGCTTTTGACCG | 1558 |
| 1907 | Db | TCTTGGAGGATGAGAAGGGGAGAGGTAAACATTTCTAATTGCAATCATAGATTTTGGAAAAA | 1966 |
| 1559 | Qy | AGTACGCGTTTTCGTAAGAGGCTTCCT-----GGCACTCACACGCGCTTCCTGTTTG | 1612 |
| 1967 | Db | CGCTAGAAGCATATGAATACATATGGCGCTGTGTGGGGCATCAACACACGTTTTCCTATGG | 2026 |
| 1613 | Qy | GACTCTCTGAACAGGAGACACAGAGGCCACTGTGAGAAAGAGTCTCTGTCTGGAAGGTCTCGC | 1672 |
| 2027 | Db | GCCTGTTAAGTATAGAGGGGAGAGAGATGAGAGAAATCTTTCACTCCGCGTGTCTC | 2086 |
| 1673 | Qy | CGCATCATAGATGGACCTGTTTTCAGTGGATCAAAAGCAAAAGCTCAGAGCGACGGTCCA | 1732 |
| 2087 | Db | AGGG-----GAGGAACCTGATGCACTGGGTCCGCTCCGTGCAGCTGC----- | 2128 |
| 1733 | Qy | CCCTGACAGAGGGCTCTTTGGAGTCTTTCAGCTGCTTTTACGAGATCCAGGAGGAGGT | 1792 |
| 2129 | Db | TGCTGACGCCACACTCTCTGGAGTCCCTCCACTGCTTGTACGAGACTTCGGAACAAACCGT | 2188 |
| 1793 | Qy | TTATCCAGACGGCCCTGAGCCACTTCCAGGTGATCGTGTTCAGCAACATTTGCCTCCAAGA | 1852 |
| 2189 | Db | TCCTGACACAAGTGATGGCCCATTTTCAAGAAA---TGGGCATGTGTGTAGAAACAGACA | 2245 |
| 1853 | Qy | TGGAGCACATGGTCTCTCTGTTCTGTTGAAGCGCTGCAGGAGCGCCACGGTCTGCACCT | 1912 |
| 2246 | Db | TGGAGCTCTTAGTGTGCACCTTCTGCAATTAATTCAGCCGCCACGTTGAAGAAGCTTCAGC | 2305 |
| 1913 | Qy | TGTTATGGCGCCACTACAGCGCGGACCGGGGAAGACCGCGGAGGTGCTCCGACGAGCGC | 1972 |
| 2306 | Db | TGATT-----GAGGGCAGGCAGCACAGATCAACATGTGAGCC | 2341 |
| 1973 | Qy | ACACGCTGTTGGTGCAGCTCAGACACAGAGAGGACCGTCTCTGTGGAGCGCTCAGTGAAC | 2032 |
| 2342 | Db | CCACCAATGGTAGTCTCTGTTTCAAG-----GTGGGTCCAGTTCACAGATGCCCTATTGGCAGA | 2395 |
| 2033 | Qy | ATCTGCGACGGCCCTGTGCACCAATCCAAACCTGATAGAGTGTCTCTGTATCCGAAATG | 2092 |
| 2396 | Db | TTCTCTTCTCCGTCCTCAAGGTCAACCAGAAAACCTGAAGGAGCTGGACCTTAAGTGGAACT | 2455 |
| 2093 | Qy | CCCTGGGACGCGGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGCNAAC | 2152 |
| 2456 | Db | CGCTGAGCCACTCTGCAGTGAAGAGTCTTTTGTAAAGACCCCTGAGACGCCCTCGCTGCCCTC | 2515 |
| 2153 | Qy | TTTCAGAACCTGAGGCTGAAGAGGTGCCGATCTCCAGCTCAGCTCGGAGGACCTCTCTG | 2212 |
| 2516 | Db | TGGAGACCTTGCCTGTTGGCTGTGGCTCTCACTGAGGACCTTCAAGGACCTTGCCT | 2575 |
| 2213 | Qy | CAGCTCTCTAGCCAAATAAGAAATTTGACAAAGATGGATCTCAGTGGCAACGGCGTTGGAT | 2272 |
| 2576 | Db | TTGGGCTGAGGCCAACCCAGACCCCTGACCGAGCTGGACCTGAGCTTCAATGTCTCACGG | 2635 |
| 2273 | Qy | TCCACGACATGATGTCTGTTTTCGAGGGCGCTCGGCCATCCCCAGTGAGGCTGCAGATGA | 2332 |
| 2636 | Db | ATGCTGGAGCCAAACACCTTTGCCAGAGACTGAGACAGCGGACCTGCAAGACTACAGCGAC | 2695 |
| 2333 | Qy | TTTCAGTTGAGGAAGTGTACGTGGATCCGGGCTGTTGAGGAGATGGCTTCTGTGCTCG | 2392 |
| 2696 | Db | TGACGTGGTTCAGCTGGCCCTCAGCTCTGACTGCTGTCAGGACCTTGGGCTCTGTGCTTA | 2755 |

Db 2516 TGGAGACCTCGGGTGTGGCTGTGGCCCTCACAGCTGAGGACTGCAAGGACCTTGCCCT 2575
Qy 2213 CAGCTCTCATAGCAATTAAGAAATTTGCAAGGATGGATCTCAGTGGCAACGGCGTTGGAT 2272
Db 2576 TTGGGCTAGAGCCCAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAGG 2635
Qy 2273 TCCAGGAGATGATGCTGTTTGGAGGGCCCTGGCGCATCCCAAGTGCAGGCTGCAGATGA 2332
Db 2636 ATGCTGGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGAC 2695
Qy 2333 TTCAAGTGAAGAGTGTGAGTGGAGTCCGGGCTTGTCAAGAGATGGCTTCTGTCTCG 2392
Db 2696 TGCAGCTGCTGAGCTGTGGCTCAGCTGACTGCTGAGCTGCTGAGGAGCTGGCCCTCTGTGCTTA 2755
Qy 2393 GCACCAACCCACATCTGCTGAGTTGAGCTTGACCTGACAGGAAATGCACTGGAGGATTTGGGCC 2452
Db 2756 GTGCCAGCCGAGCTGAGAGAGCTAGACCTGACAGCAACAACTTGATGACCTTGGCG 2815
Qy 2453 TGAGTTACTATGCCAGGAGCTGAGGACCCAGTCTGCAGACT 2495
Db 2816 TGCAGCTGCTGTGAGGGGCTCAGGCACTCTGCTGCAAACT 2858

RESULT 6

US-08-910-731-1
; Sequence 1, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHANTERJEE, DEB K.
; APPLICANT: SHANDILTA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1371 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: CDNA
; FEATURE:

; NAME/KEY: CDS
; LOCATION: 1..1368
US-08-910-731-1
Query Match 9.6%; Score 299; DB 2; Length 1371;
Best Local Similarity 54.3%; Pred. No. 1.le-69;
Matches 605; Conservative 0; Mismatches 510; Indels 0; Gaps 0;
Qy 1971 GCACACCTGTGTGTCAGCTCAGACCAGAGAGGACCGCTTCTGCTGGACGCTTACAGTGA 2030
Db 63 GCTCCAGAGTATGAGGTGTGTCAGCTCGAGACTGCGGCTCAGGAGGAGCACTGCAA 122
Qy 2031 ACATCTGGAGCGGCCCTGTGCAACCAATCAAACTCAATAGAGCTGTCTGTATACCGAAA 2090
Db 123 GGACATCGGTTTCTCCCTCCGGGCCAACCCCTCCCTGACAGAGCTGTGCTCCGACCAA 182
Qy 2091 TGGCCTGGGAGCGGGGGGTGAGCTGTCTGTCAGAGACTCAGACACCCCAACTGCAA 2150
Db 183 CGAGCTGGGCGATGCCGGCGTGCACCTGGTGTCTGAGGGCTTGAGAGCCCCCACTGCAA 242
Qy 2151 ACTTCAGAACTCGAGGCTGAAGAGTGCCTGATCTCCAGCTCAGCCTGCGAGGACCTCTC 2210
Db 243 GATCCAGAGCTCAGCTGCAAGAACTGCTCCCTGACCGAGCGGGCTGCGGGTCTCTGCC 302
Qy 2211 TGCAGCTCTCATAGCCCAATTAAGAAATTTGACAGAGATGATCTCAGTGGCAACGGCGTTGG 2270
Db 303 CAGCACGCTGCGCTCCCTGCCACGCTGCGGAGCTGCATCTCAGCGACAAACCCACTGGG 362
Qy 2271 ATTCCAGGAGATGATGCTGCTTTGCGAGGGCCCTGCGGCATCCAGTGCAGGCTGCAGAT 2330
Db 363 GGAGCGCGGCTGCGGCTGCTCTGTGAGGGGCTCCTGGACCCCCAGTGCACCTGGAGAA 422
Qy 2331 GATTCACTTGAGGAAGTGTGAGTGGAGTCCGGGGCTTGTGAGGAGATGGCTTCTGTGCT 2390
Db 423 GCTGCACTTGGAGTACTGCCGCTGACGGCGCGCAGCTGCGAGCCCTGGGCTCGGTGCT 482
Qy 2391 CGGCAACCAACATCTGTTGAGTTGAGTGCAGGAGAAATGATCTGAGAGATTTGGG 2450
Db 483 CAGGGCCACGCGGGCTTGAAGGAGCTCAGGCTGAGCAACAAACACATCGCGAGGCGG 542
Qy 2451 CTTGAGGTTACTATGCCAGGAGCTGAGGACCCAGCTGCGAGACTACGAGCTTGTGCTG 2510
Db 543 CGCCCGGGTGTGGGCTCAGGGTCTGCGAGACTCTGCTGCGAGCTGAGAGGCTCAGGCT 602
Qy 2511 GAAGATCTGCGGCTCACTGCTGCTGTGAGAGCTGAGGCTGAGCTTCAACTCTCAGTGTGA 2570
Db 603 GGAGAACTGCGGTCTCACGCCAGCAACTGCAAGAGCTGTGCGGAATTTGGGCTCCCA 662
Qy 2571 CCAGAGCTGAGAGAGCTGAGCTGAGCTGAAATGAGCTGGGGGACCTCGGGGTGCTGCT 2630
Db 663 GGCCTCGCTGAGGAGCTTGACCTGGGCGAGCAACGGCTGCGGCGAGCGGCGATAGCGGA 722
Qy 2631 GCTGTGTGAGGGCTCAGGACCTCCAGCTGCAAGCTCAGACCTCGGGTGGGCACTG 2690
Db 723 GCTGTGCCCCGGGCTCTTGAGCCCCGCTCCCGCTCAAGACCTGTGAGGCTCTGGGAGTG 782
Qy 2691 CCGGCTGGGCTGTGCGGCTGTGAGGGTCTTTCTGTGTTGTGCTCAGGCGCAACCAACCT 2750
Db 783 TGACATCACCGCCAGTGGCTGCAGAGACTCTGCCGTGTCTCCAGGCCAAGAGACCT 842
Qy 2751 CCGGGAGCTGGAATTGAGTTTCAACGACTGGGAGACTGGGGCTGTGTTGCTGGCTGA 2810
Db 843 GAAGGAGCTCAGTCTGGCGGCAACAAAGCTGGGCGAGGAGGGCGCGGCTGTGTGCGA 902
Qy 2811 GGGGCTGCACATCCCGCTGCAGACTCCAGAACTGTGGCTGATAGTGTGGCTTCAC 2870
Db 903 GAGCTGTGAGCCCGGCTGCCAGCTGGAGTCCCTGTGAGGCTGAAGTCTCTGACGCTCAC 962
Qy 2871 AGCCAAGGCTTGTGAGAAATCTTTACTTCACTTGGGGATCAACAGACCTTTCAGCGACT 2930
Db 963 GCGGCGCTGCTGCCAGCAGCTCAGCTTGAATGCTGACCAAGCAAGCATCTCTCGGAACT 1022
Qy 2931 TTACCTGACCAACAAACCCCTAGGGGACACAGGTGTCCGACTGCTTTTTCAGAGCGGCTGAG 2990

Wed Feb 9 10:22:12 2005

Db 1023 TCAGTTGAGCAGCAACAAGCTGGGTGACTCTGGCATCCAGGAGCTGTGCCAGGCCCTGAG 1082
Qy 2991 CAACTCTGGCTGCAAACTCGAGTCTCTGGTTATTGGGATGGACCTGAATAAAATGAC 3050
Db 1083 CCAGCGGGGACACACTGCGGGTGTCTGTCTTGGGAGCTGTGAGTGACCAACAGCGG 1142
Qy 3051 CCACAGTAGTGTGGCAGCGCTTCGAGTAACAAAC 3085
Db 1143 CTGACAGAGCTCGCCTCGCTCTGCTGGCCAACC 1177

RESULT 7

US-08-795-395-1
; Sequence 1, Application US/08795395
; Patent No. 5965399
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; TITLE OF INVENTION: Cloning and Expression of Rat Liver and
; TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,395
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1371 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1368
US-08-795-395-1
Query Match 9.6%; Score 299; DB 2; Length 1371;
Best Local Similarity 54.3%; Pred. No. 1.1e-69;
Matches 605; Conservative 0; Mismatches 510; Indels 0; Gaps 0;
Qy 1971 GCACACGCTGTGTGTCAGCTCAGACACAGAGAGCGCTTCTGTGGAGCGCTACAGTGA 2030
Db 63 GCTCCAGAGTATGAGTGTGTCAGGCTCGAGGCTCGGCTCAACGAGGAGCACTGCAAA 122
Qy 2031 ACATCTGGACGCGCCCTGTGCAACCAATCCAAACCTGATGAGCTCTCTGTACCGAAA 2090
Db 123 GGACATCGTTCTGCCCTCCGGGCCAACCCCTCCCTGACCGAGCTGTGCTCGCACCAA 182
Qy 2091 TGCCTGGGACGCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCAACTGCAA 2150

Db 183 CGAGCTGGGCGATGCGGGGTGCACCTGTGTGTCAGGGGCTGCGAGAGCCCACTTGC 242
Qy 2151 ACTTCAGAACTGAGGTGCGCATCTCCAGCTCAGCTCAGCTGCGAGAGACCTCTC 2210
Db 243 GATCAGAACTGAGCTGCGCATCTCCCTGACCGAGGGGCTGCGGGTCTCTGCC 302
Qy 2211 TGCAGCTCTCATAGCCAAATTAAGAAATTTGACAAGGATGATCTCAGTGGGCAACGGCTTGG 2270
Db 303 CAGCAGCTGGCTCCCTGCCACGCTGCGGAGCTGCTCTCAGCGACAACCACTGGG 362
Qy 2271 ATTCCAGAGGATGATGCTTTGCGAGGGGCTGCGGCATCCCGAGTGCAGGCTGCAGAT 2330
Db 363 GGACCGGGCTGCGGGTGTCTGTGAGGGGCTCTTGACCCCGAGTGCAGCTGCGAGTGC 422
Qy 2331 GATTGAGTTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2390
Db 423 GCTGAGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 482
Qy 2391 CGGCACCAACCCACATCTGTTGAGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2450
Db 483 CAGGGCCACGCGGGCTTGAAGGAGCTCAGCGTGAGCAACAACGACATCGCGGAGCGCG 542
Qy 2451 CCTGAGTTACTATGCCAGGGAGCTGAGCACCCAGTCTGCGAGTGTGAGTGTGAGTGTGAGTGT 2510
Db 543 CGCCCGGGTGTGCGGCGAGGTCTGGCAGATCTGCTGCTGCCAGCTGGAGAGCGCTCAGGCT 602
Qy 2511 GAAAGATCTGCGGCTCTCACTGCTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2570
Db 603 GGAGAACTGCGGTCTCTCAGCCAGCAACTGCAAAAGACCTGTGCGGAAATTTGGGCTCCCA 662
Qy 2571 CCAGAGCTGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 2630
Db 663 GCGCTGCTGAGGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 722
Qy 2631 GCTGTGTGAGGGCTCAGGCACTCCCACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 2690
Db 723 GCTGTGCGGGCTCTTGGGCGGCTTGGGCGGCTTGGGCGGCTTGGGCGGCTTGGGCGGCT 782
Qy 2691 CCGGCTGGGCTCTGCGGCTGTGAGGCTTCTTCTGTGTGCTTCCAGGCGCAACCAACCT 2750
Db 783 TGACATCACCGCAGTGGCTGAGAGACCTCTGCGGTGTCTTCCAGGCGCAACCAACCT 842
Qy 2751 CCGGAGCTGGAGTTGAGTTTCAACGACCTGGGAGAGCTGGGCTGTGTGTGTGTGTGTGTGT 2810
Db 843 GAAGAGCTCAGTCTGGCGGGCAACAAGCTGGGCGAGAGGGCGGCGGCTGTGTGTGTGTGT 902
Qy 2811 GGGGCTGCAACATCCCGCTGAGAGCTCCAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2870
Db 903 GAGCTGTCTGACGCGGGCTGCGAGTGGAGTCCCTGTGGGTGAGGCTCTGAGCTCTGAGCTCAC 962
Qy 2871 AGCCAGGCTGTGAGAAATCTTTTACTTTCACCTGGGGATCAACAGAGCTTGAACGAGCT 2930
Db 963 GCGGCGCTGTGCGAGCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1022
Qy 2931 TTACCTGACCAACAACGCGCTTGGGAGACAGGTGTCCAGCTGCTTGTGAGGGGCTGAG 2990
Db 1023 TCAGTTGAGCAGCAACAAGCTGGGTGACTCTGAGTCCAGGAGCTGTGCGAGGCTCTGAG 1082
Qy 2991 CCATCTGGCTCAAACTCCGAGTCTCTGTGTTATTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3050
Db 1083 CCAGCGGGGACCAACACTGCGGGTGTCTGTCTTGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1142
Qy 3051 CCAGTAGTGTGAGGAGCTTCCAGTAAACAAAC 3085
Db 1143 CTGACAGAGCTCGCCTCGCTCTGCTGGGCAACC 1177

RESULT 8

US-08-910-731-7
; Sequence 7, Application US/08910731
; Patent No. 5912440
; GENERAL INFORMATION:

;; APPLICANT: CHATTERJEE, DEB K.
;; APPLICANT: SHANDILYA, HARINI
;; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
;; NUMBER OF SEQUENCES: 16
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
;; CITY: WASHINGTON
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005-3934
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/910,731
;; FILING DATE: (Herewith)
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/795,395
;; FILING DATE: 04-FEB-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/794,546
;; FILING DATE: 03-FEB-1997
;; APPLICATION DATA:
;; APPLICATION NUMBER: 60/024,057
;; FILING DATE: 16-AUG-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: ESMOND, ROBERT W.
;; REGISTRATION NUMBER: 32,893
;; REFERENCE/DOCKET NUMBER: 0942.3440003
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-371-2600
;; TELEFAX: 202-371-2540
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1371 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: both
;; TOPOLOGY: both
;; MOLECULE TYPE: CDNA
US-08-910-731-7

Query Match 9.4%; Score 293.2; DB 2; Length 1371;
Best Local Similarity 54.2%; Pred. No. 3.8e-68;
Matches 595; Conservative 0; Mismatches 503; Indels 0; Gaps 0;
Qy 1971 GCACACGCTGTTGGTGCAGCTCAGACCCAGAGAGGACCGTTCTGCTGGACGCTTACAGTGA 2030
Db 63 GCTCCAGCAGTATGAGGTGGTTCAGGCTCGACGCTGCGGCTCAGGAGGAGCACTGCA 122
Qy 2031 ACATCTGGCAGCGGCTGTGCAACCAATCCAACTGATAGAGCTGTCTGTACCGAAA 2090
Db 123 GGACATCGGTTCTGCTCCGCGGCAACCCCTCCCTGACCGAGCTGTGCTTCGACCAA 182
Qy 2091 TGCCTTGGCAGCGGCGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGCA 2150
Db 183 CGAGCTGGCGATCGCGCGGTGCACTGTGCTGCTGAGGCTCGAGAGCCCACTGCA 242
Qy 2151 ACTTCAGAACCTGAGGCTGAAGAGGTGCGGCTCTCCAGCTTCAGCTGCGAGGACCTCTC 2210
Db 243 GATCCAGAAGCTCAGCTGCGAGAACTGCTCCCTGACCGAGCGGCTGCGGGTCTGCTGCC 302
Qy 2211 TGCAGCTCTATAGCCAAATAGAAATTTGAAGAAGATGATCTAGTGGCAAGCGGCTTGG 2270
Db 303 CAGCACGCTGGCTCCCTGCGCCACGCTGCGGGAGCTGCACTCAGCGACACCACTGGG 362
Qy 2271 ATTTCCAGGATGATGCTGCTTTTCGAGGGCTCGCGGCTCCAGTCCCGAGGCTGCAGAT 2330
Db 363 GGACCGCGGCTGCGGCTGCTCTGTAGGGGCTCTTGACCCCGAGTGCACCTGGAGAA 422

Qy 2331 GATTCACTTGAGGAAGTGTGAGCTGAGTCCGGGGCTTGTTCAGGAGATGGCTTCTGTGCT 2390
Db 423 GCTGCACTTGAGTACTGCGGCTTGAAGGAGCTGCGGCGCCAGCTGCGAGCCCTTGGCTGCT 482
Qy 2391 CGGCACCAACCCACATCTGTTGAGTTGAGCTGACAGGAATGCACCTGGAGGATTTTGGG 2450
Db 483 CAGGGCCACCGGGGCTTGAAGGAGCTCAAGTGAACCAACAGACATCGGAGGCGCGG 542
Qy 2451 CCTGAGGTTACTATGCCAGGAGCTGAGGACCCAGTCTGCAGACTACGAGATTTTGTGCT 2510
Db 543 CGCCCGGCTGCTGGGCGCAGGCTTGGCCGACTCTGCTGCCAGCTGGAGACGCTCAGGCT 602
Qy 2511 GAAGATCTGCGGCTCACTGCTGCTGTGCTGTGAGAGTGGCCCTCAACTCTCAGTGTGA 2570
Db 603 GGAGAACTGCGGCTCTCAGCGCAGCAACTGCAAGAGCTGTGCGGAATTTGTGGCCTCCA 662
Qy 2571 CCAGAGCTGAGAGAGCTGGACCTGAGCTGAATGAGCTGGGAGCTCGGGGTGCTGCT 2630
Db 663 GGCCTCGCTGCGGAGCTGGCCCTGGGAGCAAGAGCTGGGTGATGTGGGCATGGCGGA 722
Qy 2631 GCTGTGAGGGGCTCAGGAGCTCCAGCTGCAAGCTCCAGACCTCGGCTGGGCTGGGCTCTG 2690
Db 723 GCTGTGCCAGGCTGCTCCACCCAGCTCCAGGCTCAGGACCTGTGGATCTGGGAGTG 782
Qy 2691 CCGGCTGGGCTTGCCTGCTGTGAGGCTTCTTCTGTGTGCTCCAGGCAACCAACCACT 2750
Db 783 TGGCATCACTGCCAAGGGCTGCGGGATCTGTGCCGTGTCTCAGGGCCAGAGAGCT 842
Qy 2751 CCGGAGCTGGACTTGTGAGTTTCAACGACCTGGGAGCTTGGGCTGTGGTGTGCTGGCTGA 2810
Db 843 GAAGGAGCTCAGCTGCGCGCAACGAGCTGGGGGATGAGGGTCCCGACTGCTGTGTGA 902
Qy 2811 GGGGCTGCAACATCCCGCTCGACACTCCAGAACTGTGGCTGGATAGCTGTGGCTCAC 2870
Db 903 GACCTGCTGGAACCTGGCTGCCAGCTGGAGTGCCTGTGGGTGAAGTCTTGCAGCTTCA 962
Qy 2871 AGCCAAGCTTGTGAGAACTCTTTTACCTCCCTGGGGATCAACAGACCTTTCAGCGACT 2930
Db 963 AGCGCTGCTGCGCCCACTTCAGCTCAGTGTGCTGCGGCGCAGAACAGGTTTCTCTCTG 1022
Qy 2931 TTACTTCACCAACACCGCTAGGGGACAGGTTCCGAGCTGCTTGGATGACCTGTAATAAATG 2990
Db 1023 ACAGATAAGCAACACAGGCTGGAGGATGCGGGCTGCGGAGCTGTGCCAGGCTGGG 1082
Qy 2991 CCATCTGGCTGCAAACTCCGAGTCTCTGTTTATTTGGATGACCTGTAATAAATGAC 3050
Db 1083 CCAGCTGCTGCTGTGCTGCGGGTGTCTGTGGTGGCGACTGCGCATGTGATGACAGCAG 1142
Qy 3051 CCACAGTAGTTGGCAGC 3068
Db 1143 CTGCAGCAGCTGCGCGC 1160

RESULT 9

US-08-910-731-5
; Sequence 5, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

| | | | |
|------|----|---|------|
| 2586 | QY | GCTGCACTCGAGCTGAATGAGCTGGGGGACCTCGGGGTGCTGCTGCTGCTGAGGGCCT | 2644 |
| | | | |
| 864 | Db | GCTCAGCCTGGCGGCAACGAGCTGGGGGATGAGGGTGCCCGACCTGCTGTGTGAGACCT | 923 |
| | | | |
| 2646 | QY | CAGGATCCACGTCGTCAGCTCCAGACCTCGGGTGGGCATCTGCCGGCTGGGCTCTGC | 2705 |
| | | | |
| 924 | Db | GCTGGAACCTGGCTGCAGCTGAGTCTGTGTGGGTGAAGTCTCTCGAGCTTCACAGCGC | 983 |
| | | | |
| 2706 | QY | CGCCTGTGAGGGTCTTTCTGTGGTGTCTCAGCTCAGTGTGCGCCAGAACAGGTTTCT | 2765 |
| | | | |
| 984 | Db | CTGCTGCCCCCACTTCAGCTCAGTGTGCGCCAGAACAGGTTTCTCTCGAGCTTCACAGCGC | 1043 |
| | | | |
| 2766 | QY | GAGTTTCAACGACTGGGAGACTGGGSCCTGTGTTGCTGTGGCTGAGGGCTGCAACATCC | 2825 |
| | | | |
| 1044 | Db | AAGCAACACAGGCTTGAGGATGCGGGCTGTGCGGAGCTGTGCCAGGGCTTGGGCGCAGCC | 1103 |
| | | | |
| 2826 | QY | CGCCTGCAGACTCCAGAAACTGTGGCTGGATAGCTGTGGCTCACAGCCTCAAGGCTTGTGA | 2885 |
| | | | |
| 1104 | Db | TGGCTGTGCTGGGGTGCTCTGTTGGCCGACTGCGATGTGAGTGACAGCAGCTGCAG | 1163 |
| | | | |
| 2886 | QY | GAATCTTTACTTACCTGGGATCAACAGACCTTGACCGACTTTACCTGACCAACAA | 2945 |
| | | | |
| 1164 | Db | CAGCCTCGCGCAACCTGTTGGCCACACACAGCTGCGTGAGCTGAGCTGACCTCAGCAACAA | 1223 |
| | | | |
| 2946 | QY | CGCCTAGGGGACACAGGCTGCCACTGCTTTGCAAGCGGTGAGCCTCTCTGCTGCA | 3005 |
| | | | |
| 1224 | Db | CTGCTGGGGACGCGGCATCTCTGACGTGTGTGAGAGCGTCCGGCAGCCGGGCTGCCT | 1283 |
| | | | |
| 3006 | QY | ACT 3008 | |
| | | | |
| 1284 | Db | CCT 1286 | |

RESULT 10
 US-08-910-731-3
 ; Sequence 3, Application US/08910731
 ; Patent No. 5932440
 ; GENERAL INFORMATION:
 ; APPLICANT: CHATTERJEE, DEB K.
 ; APPLICANT: SHANDILYA, HARINI
 ; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 ; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
 ; CITY: WASHINGTON
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3934
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/910,731
 ; FILING DATE: (Herewith)
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/795,395
 ; FILING DATE: 04-FEB-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/794,546
 ; FILING DATE: 03-FEB-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/024,057
 ; FILING DATE: 16-AUG-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ESMOND, ROBERT W.
 ; REGISTRATION NUMBER: 32,893
 ; REFERENCE/DOCKET NUMBER: 0942.3440003
 ; TELECOMMUNICATION INFORMATION:

| QY | 2826 | CGCCTGCAGACTCCAGAACTGTGGCTGGATAGTGTGGCTTACAGCCAAAGGCTTGTA | 2888 |
|--|------|---|------|
| Db | 1089 | GGACACAGTGTGGCTGTGGCTTTGGCTGGGAGACTGTGATGTGACAGACAGTGGCTGCAG | 1148 |
| QY | 2886 | GAATCTTTTACTTCAACCTGGGGATCAACAGACCTTACCGACCTTTACCTGACCAACAA | 2945 |
| Db | 1149 | CAGCCTTGCCACTGTCTCTGGCCAAACCGCAGCTTGAGGAACTGGACCTCAGTAACAA | 1208 |
| QY | 2946 | CGCCCTAGGGGACACAGGTGTCCGACTGCTTTTGAAGCGGCTGAGCCATCTCTGGCTGCAA | 3005 |
| Db | 1209 | CTGATGGGGACAAACGGTGTCTTACAACTGCTGGAGAGCTCAACAGCCAGCTGCAT | 1268 |
| QY | 3006 | ACTCGAGTCTCTGGTTATTTGGGATGGACCTGA | 3040 |
| Db | 1269 | CCTTCAGCAGCTTGTCTGTATGACATTTACTGGA | 1303 |
| <p>RESULT 11</p> <p>US-08-795-395-3</p> <p>Sequence 3, Application US/08795395</p> <p>Patent No. 5965399</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: CHATTERJEE, DEB K.</p> <p>APPLICANT: SHANDILYA, HARINI</p> <p>TITLE OF INVENTION: Cloning and Expression of Rat Liver and</p> <p>TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor</p> <p>NUMBER OF SEQUENCES: 8</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.</p> <p>STREET: 1100 NEW YORK AVE., N.W., SUITE 600</p> <p>CITY: WASHINGTON</p> <p>STATE: D.C.</p> <p>COUNTRY: USA</p> <p>ZIP: 20005-3934</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: PatentIn Release #1.0, Version #1.30</p> <p>CURRENT APPLICATION DATA: US/08/795,395</p> <p>FILING DATE: 04-FEB-1997</p> <p>CLASSIFICATION: 435</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: 60/024,057</p> <p>FILING DATE: 16-AUG-1996</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: ESMOND, ROBERT W.</p> <p>REGISTRATION NUMBER: 32,893</p> <p>REFERENCE/DOCKET NUMBER: 0942.3440002</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: 202-371-2600</p> <p>TELEFAX: 202-371-2540</p> <p>INFORMATION FOR SEQ ID NO: 3:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 1374 base pairs</p> <p>TYPE: nucleic acid</p> <p>STRANDEDNESS: both</p> <p>TOPOLOGY: both</p> <p>MOLECULE TYPE: cDNA</p> <p>FEATURE:</p> <p>NAME/KEY: CDS</p> <p>LOCATION: 1..1368</p> <p>US-08-795-395-3</p> | | | |
| <p>Query Match</p> <p>Best Local Similarity 8.3%; Score 256.6; DB 2; Length 1374;</p> <p>Matches 556; Conservative 0; Mismatches 499; Indels 0; Gaps 0;</p> | | | |
| QY | 1986 | GCAGCTCAGACAGAGAGACCGTTCTGTGGACCCCTACAGTGAACATCTGGCAGCGGC | 2045 |
| Db | 249 | GAAGCTGAGCCTTCAGAACTCAGAGTGTGACGAAGCTGGCTGTGGGTCCTGCCTGATGT | 308 |
| QY | 2046 | CCTGTGCACCAATCCAAACCTCATAGAGCTGTCTGTACCCGAATGCCCTGGCGACCGC | 2105 |
| Db | 309 | GCTGCCCTTTTGTCTACCTCGGTGAATCATCTCAATGACAACTCTGGGGATGA | 368 |
| QY | 2106 | GGGGGTGAAGTCTCTGTCAAGGACTCAGACACCCCAACTGCAACTTCAGAACTTGAG | 2165 |
| Db | 369 | AGCCTGAAGTCTCTGTGAAGGACTCGGGACCCCGAGTGGCTCTTGAGAAGCTTCA | 428 |
| QY | 2166 | GCTGAAGAGTGGCGCATCTCAGCTCAGCTCAGCTGAGGAGCTCTCTGAGCTCTCATAGC | 2225 |
| Db | 429 | GTTGGAATACGTAACTCACAGCTACAGCTGCGAGCCCTGGCCTCAGTGTCTCAGGT | 488 |
| QY | 2226 | CAATAGAAATTTACAAAGATGGATCTCAGTGGCAACGGCTTGGATTCACAGCATGAT | 2285 |
| Db | 489 | GAACCTGACTTTAAAGACTAGTANTGAGCAACAATGACTTCCATGAGGCTGGTATCCA | 548 |
| QY | 2286 | GCTGCTTTGCGAGGGCTCGCGATCCCGAGTGCAGGCTGCAGATGATTCAGTTGAGGAA | 2345 |
| Db | 549 | CACCTGTGCGAGGGCTGAAGATTTGCTGTCTCACTGGAGTCACTCAAACTGGAGAA | 608 |
| QY | 2346 | GTGTGAGTGGAGTCCGGGGCTTGTGAGAGATGCTTCTGTGTGGCCACCAACCCACA | 2405 |
| Db | 609 | CTGTGTATCACATCAGCAACCTGCAAGGATCTGTGTGATGTGTGGCTCCAAAGCCTC | 668 |
| QY | 2406 | TCTGTTGAGTGGACCTCAGAGGAATGCACTGAGGATTTGGCCCTGAGTTACTATG | 2465 |
| Db | 669 | ACTGCAAGAACTGGACTTGGGAGCAACAAGCTGGGCAACACAGGCAATTCAGCACTGTG | 728 |
| QY | 2466 | CCAGGACTGAGGCACCCAGTCTGCAGACTACGGAATTTGTGGCTGAAGATCTGCCGCT | 2525 |
| Db | 729 | CTCAGAGCTGCTTCCAGCTGAGGCTGAGGACTCTGTGGCTCTGGGACTGTGATGT | 788 |
| QY | 2526 | CAGTGTGCTGTGTGACAGCTGGCCCTCAACTCTCAGTGTGAACACAGACCTTGAGGA | 2585 |
| Db | 789 | CAGTGCAGAGGCTGCAAGGACCTGTGCCGTCTCTCAGAGCAAGCAGAGACCTTGAAGGA | 848 |
| QY | 2586 | GCTGGACCTGAGCTGAATGAGCTGGGGACCTCGGGTGTGCTGTGTGAGGGCCT | 2645 |
| Db | 849 | ACTCAGCCTAGCTGGCAATGAGCTGAAGGATGAGGGTCCCAACTGCTGTGTGAGAGCCT | 908 |
| QY | 2646 | CAGGCATCCAGCTGCAAGCTCCAGACCTCGGGTGGGCATCTGCCGCTGGGCTCTGC | 2705 |
| Db | 909 | GTTAGAGCTTGGCTGTACCTGGAGTCACTGTGGGTAAAGACCTGTAGCCTTCACAGCTGC | 968 |
| QY | 2706 | CGCCTGTGAGGGTCTTTCTGTGGTGTCTCCAGGCCAACCAACCTCCGGGAGCTGGACTT | 2765 |
| Db | 969 | CTCTGTCCCACTTCTGCTCGTGTGTGACCAAAAACAGATTTCTGTGTGAGTTGCAAT | 1028 |
| QY | 2766 | GAGTTTCAACAGACTGGGAGACTGGGGCTGTGTGTGTGCTGAGGGGCTGCAACATCC | 2825 |
| Db | 1029 | GACAGCAACCCGCTGGGAGACTCGGAGTCTGTGAGCTTTGCAAGGCCCTGGGCTATCC | 1088 |

| | | | |
|----|------|---|------|
| QY | 2046 | CCTGTGCACCAATCCAAACCTGATAGAGTGTCTCTGTATCCGAAATGCCCCTGGCAGCGG | 2105 |
| DB | 309 | GCCTGGCTCTTTTGTCTACCTCGGTGAACTACATCTCAATGACAAACCTCTGGGGATGA | 368 |
| QY | 2106 | GGGGGTGAAGTCTCTGTCAAGAGCTCAGACACCCCACTGCAACTTTCAGAACTTGAG | 2165 |
| DB | 369 | AGGCCTGAAGTCTCTGTGAAGGATCCCGGGACCCCGAGTCGCTTTTGAAGGCTTCA | 428 |
| QY | 2166 | GCCTGAAGAGTGGCGGATCTCCAGCTCAGCCTCGCAGGACCTCTCTGCAGCTCTCATAGC | 2225 |
| DB | 429 | GTTTGGAAATCTGTAACTCAAGCTACAGCTCGAGCCCTGGCTCAGTGTCTCAGGCT | 488 |
| QY | 2226 | CAATAAGAAATTTGACAAGAGTGGATCTCAGTGGCAACGGGTGGAATTCAGGCAATGAT | 2285 |
| DB | 489 | GAACACCTGACTTTAAAGAGCTAGTATTGAGCAACAATGACTTCATGAGGCTGGTATCCA | 548 |
| QY | 2286 | GCTGCTTTGGAGGGCTCGGGATCCCAAGTCGAGCTGCAGATGATTCAGTTGAGGAA | 2345 |
| DB | 549 | CACCTCTGTGCCAGGGCTGAAGAAATCTGCTGTCAACTGGAGTCACTCAAACTGGAGAA | 608 |
| QY | 2346 | GTGTCACTGGAGTCCGGGCTTTGTCAAGGATGGCTTCTGTGCTGGCAACCAACCCACA | 2405 |
| DB | 609 | CTGTGGTATCACATCAGCACTGCAAGGATCTGTGTGATGTTGTGGCTCTCAAAGCCTC | 668 |
| QY | 2406 | TCTGGTTGAGTTGGACCTGACAGAAATGCACCTGGAGGATTTGGGCTGAGGTTACTATG | 2465 |
| DB | 669 | ACTGCAAGAACTTGACTTTGGGACCAAAAGCTGGGCAACACAGGCAATTCAGCACTGTG | 728 |
| QY | 2466 | CCAGGACTGAGGCAACCCAGTCTGACAGCTAAGCACTTTGTGGCTGAAGATCTGCGCCT | 2525 |
| DB | 729 | CTCAGGACTGTGCTTCCAGCTCAGGCTGAGGACTCTGTGGCTCTGGGACTGTGATGT | 788 |
| QY | 2526 | CACCTGTGCTGCTGTGACGAGCTGGCTCAACTCTCAGTGTGAACACAGAGCCTTGAGAG | 2585 |
| DB | 789 | CACCTGCAAGAGGCTGCAAGGACCTGTGCGGTCTCTCAGAGCCAGCAGAGCTGAAGGA | 848 |
| QY | 2586 | GCTGGACCTGAGCCTGATAGCTGGGGGACCTCGGGGTGTGCTGTGTGTGAGGGCCT | 2645 |
| DB | 849 | ACTCAGCCTAGCTGGCAATGAGCTGAAGATGAGGGTGCCCAACTGTGTGTGTGAGAGCCT | 908 |
| QY | 2646 | CAGGCATCCACGTTGCAAGCTCCAGACCTCGGTTGGGCACTCTGCGGCTGGGCTCTGC | 2705 |
| DB | 909 | GTTAGAGCCTGGCTGTCACTGAGTGGAGTCACTGTGGGTAAAGACCTGTAGCTTCAAGTGC | 968 |
| QY | 2706 | CGCCTGTGAGGGTCTTTCTGTGGTGTCTCAGGCGCAACCAAACTCCGGGAGCTGGACTT | 2765 |
| DB | 969 | CTCTGTGCCCACTCTCTCTCGTGTGTGACAAAAACAGTTCTCTGTGTGAGTTGCAAT | 1028 |
| QY | 2766 | GAGTTTCAAAGACCTGGAGAGCTGGGGCTGTGTTGTGCTGAGGGCTGCAACATCC | 2825 |
| DB | 1029 | GAGCAGCAACCGCTGGGAGACTCGGGAGTCTGGAGCTTTTGAAGGCCCTGGGCTATCC | 1088 |
| QY | 2826 | CGCCTGCAGACTCCAGAAACTGTGGCTGGATGCTGGCTCTCAGCAGCAAGGCTTGGA | 2885 |
| DB | 1089 | GGACACAGTGTGCGTGTGCTTTGGCTGGAGACTGTGATGTGACAGACAGTGGCTGCAG | 1148 |
| QY | 2886 | GAATCTTTTCAACCTGGGGATCAACGAGACCTTGACCACTTTACCTGACCAACAA | 2945 |
| DB | 1149 | CAGCCTTGCACTGCTCTGCTGGCAACCGCAGCTTTGAGGAACTTGGACCTCAGTAAACA | 1208 |
| QY | 2946 | CGCCCTAGGGGACACAGGTGTCCGACTGTTTGGCAAGGGGTGAGCCATCTCTGGCTGCAA | 3005 |
| DB | 1209 | CTGCAATGGGGACACGGTGTCTCAACTGTCTGGAGAGCCTCAAAACAGCCACGCTGCAT | 1268 |
| QY | 3006 | ACTCGAGTCTCTGGTTATTTGGGATGGAACCTGA | 3040 |
| DB | 1269 | CCTTCAGCAGCTTGCTGTATGCAATTTACTGGA | 1303 |

Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 208:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SINTBST01
CLONE: 1499408
US-09-016-434-208

Query Match 2.8%; Score 85.8; DB 4; Length 257;
Best Local Similarity 60.5%; Pred. No. 5.7e-13;
Matches 155; Conservative 0; Mismatches 100; Indels 1; Gaps 1;

| | | | |
|----|------|--|------|
| Qy | 1036 | TACTTCTACAGATATTTCCAAATGCAGACGAGCGGGCCAACTCTTCAATTAGCTGAGG | 1099 |
| Db | 2 | TATTTCTACAAGATATTTCCGGGATGAGAGGAGCGCGCTACCCGCTTCGTGAG | 61 |
| Qy | 1096 | GACAAAGAGCCTCTCTTACCAATGTGTTTCGTCCTCCCTCGTGTGTGTGTGGTGGTGTGTAC | 1155 |
| Db | 62 | GAGAAAGAGACGTGTTTCGGCTGTGCTTCGTGCTCCCTTCGTGTGTGGATCGTGTGCACC | 121 |
| Qy | 1156 | TGCCTCCAGCAGCTGAGAGGTGGGGGCTGTTTGAGACAGAGCTCCAGGACCAACCACT | 1215 |
| Db | 122 | GTGCTGC-GCCANANTGNGAGCTCGGTTCGGGACCTGTTCGGCGACGTCCAAGACCAACACG | 180 |
| Qy | 1216 | GCAGTGTACATGCTTACCTGCTGAGTCTGATGCAACCAGCGGGGGGCCCGCCGCTC | 1275 |
| Db | 181 | TCAGTGTACTGTTCTTTTCATCACAGCGGTCTTGAGCTCGGCTCGGTAGCGAGCGGGCCC | 240 |
| Qy | 1276 | CAGCCCCCAACCCAAAC | 1291 |
| Db | 241 | CGGTTGACGGGCGCACC | 256 |

RESULT 13
US-09-023-655-395
; Sequence 395, Application US/09023655

RESULT 12
US-09-016-434-208
; Sequence 208. Application US/09016434

RESULT 14
US-09-016-434-1011

Wed Feb 9 10:22:12 2005

1420 CGCACCAGATTCAACCTCAAGGGCTTCTCTGAACAGGGGCAATCGAGTGTACTTGAGGAAG 1479
1048 TATTTCCACAATGACAGAGCAGCGGGGCAAGTCTTCAATTACGTGAGGGGACAACGAGCCT 1107
1480 CGCCATCATGAGCCCGGGGTGGGGACCCGCTCATCGCCCTGCTCCAGAGACCTCAGCC 1539
1108 CTCTTCCACCATGTCTTCGTCCTCCCTTGGTGTGCTGGGTGTGTACTCTGCTTCCAGCAG 1167
1540 CTGACCGGTTTGTGCCACCTGCTGCTTCTCATGGATGCTGTCCAAATGCCACAGGAA 1599
1168 CAGCTGGAGGGTGGGGGCTGTTGAGACAGACGTCCA 1204
1600 CTGTTGCTGCAGGAGGGGGTCCCAAGACCACTA 1636

Search completed: February 8, 2005, 22:51:52
Job time : 536 secs

1469 TCAGGAATT-CTTTGAGCTATGTACTATATCTCGACGAGGGGAGG 1516
423 TCNAGGAATCTCCGGGCACTGTCTTACCTCTGGAAGACGGGGG 471

RESULT 15

US-09-949-016-3209
; Sequence 3209, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3209
; LENGTH: 4485
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3209

Query Match 2.2%; Score 68.2; DB 4; Length 4485;
Best Local Similarity 48.1%; Pred. No. 1.3e-07;
Matches 364; Conservative 0; Mismatches 368; Indels 25; Gaps 5;
QY 461 GGGGACAGCGAGGACCCGCGAGCCACCGGCGCTGATGCC-CCATCAAGATAGAGACCCCTCTTT 519
DB 892 GTGGCATGTGGTGGACCCCGGAGAGCCAGCCACCTTGGGCCCTGGAGGAGCTTTC 951
QY 520 GAGCCAGAGGAGGAGCGCCGCGAGCCACCGGCGACCGTGTGTCATCAAGGGCGGCGAGG 579
DB 952 AGCAACCCCTGGCCACCTCAATGACGATCGGACACTGTGTGTGGTGGTGGAGCGGCG 1011
QY 580 ATAGCAAGTCCATGCTGGGACACAAGGTGATGTGACTGGGCGGCGGAGGAGCTCTTC 639
DB 1012 AGTGGCAAGAGCACGCTCCTGCGAGCGGCTGCATTTGTGTGGGTGCGAGGGCAAGACTTC 1071
QY 640 CAAGGAGATTGATTATCTCTTCTACATCAACTGCGAGGAGATGAACAGAGTGCCACG 699
DB 1072 CAGG---AATTTCTCTTTGTCTTCCANTCAGTCCCGGAGCTGCACTGATGCCCAA 1128
QY 700 GAATGAGCATGAAGACCTCATCTTCAAGCTGTGCGCTGAGCCCGCGGCTTCCAG 759
DB 1129 CCACTCTCTGTGGGACTCTACTCTTGAGCACTGTGTGGCTGTGATGTTGGTCAAGAA 1188
QY 760 GAGCTCATCGAGTTCCC-----GAGCGCTCTTTTCATCATGACGGCTTC 807
DB 1189 GACATCTCCAGTTACTCTTGAACCAACCTGACCGGTCTTGTAACTTTGATGGCTTT 1248
QY 808 GATGAGCTCAAGCTTTCTTTCCACGATCCTCAGGAGCCCTGTGTGCTCTGTGGGAGGAG 867
DB 1249 GAGAGTTCAAGTTCAAGTTCAAGGATCGGAACGCACTG-----CTCCCGGACCGAC 1302
QY 868 AAACGGCCCAAGGAGTGTCTTTAAAGCTTAATTCGAAGAGTGTCTCCCTCAGCTA 927
DB 1303 CCCACCTCTGTCCAGACCCCTGCTCTTCAACCTTCTGAGGGCAACCTGCTGAAGATGCC 1362
QY 928 TCTTTGCTCATACCAACAGCGCCACCGCTTTGGAGAGCTCCACCGTCTGCTGGAGCAC 987
DB 1363 CGCAAGGTGTGACCGAGCGTCCGGCGGCTGTGTCCGCTTCTCAG---GAAGTACATC 1419
QY 988 CCCAGGCAATGAGATCTCTGGGCTTCTGTGAGGCAAGAGGAGATATCTTCTACAG 1047

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2005, 20:14:19 ; Search time 1555 Seconds
(without alignments)
11504.081 Million cell updates/sec

Title: US-10-781-294-23
Perfect score: 3108
Sequence: 1 atgtctacgaaccgcagcag.....atttgacattgctgtga 3108

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4313806 seqs, 2877871033 residues

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications NA:*
- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
 - 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
 - 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
 - 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
 - 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
 - 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
 - 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
 - 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
 - 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
 - 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
 - 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
 - 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
 - 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
 - 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
 - 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
 - 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
 - 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
 - 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
 - 19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
 - 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
 - 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
 - 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 3108 | 100.0 | 3108 | 10 | US-09-965-621-23 |
| 2 | 3108 | 100.0 | 3108 | 17 | US-10-407-866-23 |
| 3 | 3108 | 100.0 | 3108 | 18 | US-10-781-294-23 |
| 4 | 2997.8 | 96.5 | 3186 | 14 | US-10-124-498-17 |
| 5 | 2997.8 | 96.5 | 3186 | 14 | US-10-066-521-17 |
| 6 | 2929 | 94.2 | 3218 | 17 | US-10-407-866-67 |
| 7 | 2644.2 | 85.1 | 3466 | 17 | US-10-108-260A-718 |
| 8 | 1705.4 | 54.9 | 1800 | 17 | US-10-162-335-21 |
| 9 | 1681.8 | 54.1 | 2158 | 18 | US-10-467-397-19 |
| c 10 | 1680 | 54.1 | 1704 | 16 | US-10-029-386-24660 |
| 11 | 1616 | 52.0 | 4931 | 15 | US-10-028-374-1 |
| | | | | | Sequence 23, Appl |
| | | | | | Sequence 23, Appl |
| | | | | | Sequence 23, Appl |
| | | | | | Sequence 17, Appl |
| | | | | | Sequence 67, Appl |
| | | | | | Sequence 718, Appl |
| | | | | | Sequence 21, Appl |
| | | | | | Sequence 19, Appl |
| | | | | | Sequence 24660, A |
| | | | | | Sequence 1, Appl |

| | | | | | | |
|------|--------|------|------|----|---------------------|-------------------|
| 12 | 1616 | 52.0 | 4931 | 16 | US-10-183-770-1 | Sequence 1, Appl |
| 13 | 1394.8 | 44.9 | 1683 | 17 | US-10-162-335-23 | Sequence 23, Appl |
| 14 | 947.6 | 30.5 | 2682 | 9 | US-09-799-983-1 | Sequence 1, Appl |
| 15 | 800.4 | 25.8 | 3102 | 13 | US-10-127-516-6 | Sequence 6, Appl |
| 16 | 800.4 | 25.8 | 3102 | 13 | US-10-027-629-6 | Sequence 6, Appl |
| 17 | 800.4 | 25.8 | 3102 | 16 | US-10-132-967-6 | Sequence 6, Appl |
| 18 | 800.4 | 25.8 | 3857 | 13 | US-10-127-516-4 | Sequence 4, Appl |
| 19 | 800.4 | 25.8 | 3857 | 13 | US-10-027-629-4 | Sequence 4, Appl |
| 20 | 800.4 | 25.8 | 3857 | 16 | US-10-132-967-4 | Sequence 4, Appl |
| 21 | 800 | 25.7 | 4193 | 17 | US-10-264-958B-21 | Sequence 21, Appl |
| 22 | 691 | 22.2 | 2511 | 9 | US-09-799-983-5 | Sequence 5, Appl |
| 23 | 639 | 20.6 | 2524 | 10 | US-09-965-621-58 | Sequence 58, Appl |
| 24 | 639 | 20.6 | 2524 | 17 | US-10-407-866-58 | Sequence 58, Appl |
| 25 | 639 | 20.6 | 2524 | 18 | US-10-781-294-58 | Sequence 58, Appl |
| 26 | 595.6 | 19.2 | 2494 | 9 | US-09-799-983-3 | Sequence 3, Appl |
| c 27 | 509 | 16.4 | 509 | 16 | US-10-029-386-10957 | Sequence 10957, A |
| 28 | 475.4 | 15.3 | 3521 | 18 | US-10-794-342-2 | Sequence 2, Appl |
| 29 | 453.4 | 14.6 | 487 | 10 | US-09-918-995-23437 | Sequence 23437, A |
| 30 | 435.6 | 14.0 | 2775 | 17 | US-10-357-820-51 | Sequence 51, Appl |
| 31 | 404.4 | 13.0 | 4035 | 14 | US-10-124-498-5 | Sequence 5, Appl |
| 32 | 404.4 | 13.0 | 4035 | 14 | US-10-066-521-5 | Sequence 5, Appl |
| 33 | 402.8 | 13.0 | 3226 | 17 | US-10-092-900A-347 | Sequence 347, App |
| 34 | 402.8 | 13.0 | 3489 | 17 | US-10-416-642-3 | Sequence 3, Appl |
| 35 | 402.8 | 13.0 | 3830 | 15 | US-10-216-645-3 | Sequence 3, Appl |
| 36 | 402.8 | 13.0 | 3926 | 15 | US-10-216-645-1 | Sequence 1, Appl |
| 37 | 401.2 | 12.9 | 3885 | 18 | US-10-860-761-3 | Sequence 3, Appl |
| 38 | 401.2 | 12.9 | 3900 | 17 | US-10-399-443-23 | Sequence 23, Appl |
| 39 | 401.2 | 12.9 | 3900 | 17 | US-10-677-943-23 | Sequence 23, Appl |
| 40 | 398 | 12.8 | 4287 | 9 | US-09-996-617-5 | Sequence 5, Appl |
| 41 | 398 | 12.8 | 4287 | 9 | US-09-931-071-5 | Sequence 5, Appl |
| 42 | 398 | 12.8 | 4422 | 9 | US-09-388-221-1 | Sequence 1, Appl |
| 43 | 398 | 12.8 | 4422 | 18 | US-10-828-920-1 | Sequence 1, Appl |
| 44 | 398 | 12.8 | 4556 | 9 | US-09-388-221-9 | Sequence 9, Appl |
| 45 | 398 | 12.8 | 4556 | 18 | US-10-828-920-9 | Sequence 9, Appl |

ALIGNMENTS

RESULT 1
; Sequence 23, Application US/09965621
; Publication No. US20030077699A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Ioredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/09/965,621
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671,760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 3108
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3105)
US-09-965-621-23

Query Match 100.0%; Score 3108; DB 10; Length 3108;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|------|---|------|
| QY | 1 | ATGCTACGAACCGCAGGCGAGCGGCTCTGTGCGCTGTCCACCTACTTTGGGAAGAACTC | 60 |
| DB | 1 | | |
| QY | 61 | GAGGCTGTGGAACCTGAAGAGTTCAAGTTATACCTGGGACCGCAGACAGAGCTGGGAGAA | 120 |
| DB | 61 | | |
| QY | 121 | GGCAAGATCCCCTGGGGAAGCATGAGAAAGCCGGTCCCTGTGAAATGGCCAGCTGCTC | 180 |
| DB | 121 | | |
| QY | 181 | ATCACCACCTTCGGGCCAGAGAGGCGCTGGAGGTTGGCTCTCAGACACTTTGAGCGGATA | 240 |
| DB | 181 | | |
| QY | 241 | AAcAGAAGGAACTGTGGGAGAGAGGACACAGAGAGGAGGACCTGGTAGGATCCCCAGGAA | 300 |
| DB | 241 | | |
| QY | 301 | ACCTACAGGAGCTATGTCGCGCAGGAAATTCGGGCTCATGGAAGACCGCAATGCGCGCTA | 360 |
| DB | 301 | | |
| QY | 361 | GGGGAATGTTCAACTCAGCACCAGGTTACACCGGCTCCTGCTGGTGAAGAGCACTCA | 420 |
| DB | 361 | | |
| QY | 421 | AAcCCCATGAGGTCAGCAGCAGCTTCTGGACACAGGCGGGGACACGCGAGGACCGTG | 480 |
| DB | 421 | | |
| QY | 481 | GGACACAGGCTAGCCCCATCAAGATAGAGACCCCTTTTGAGCCACAGGAGGCGCCCC | 540 |
| DB | 481 | | |
| QY | 541 | GAGCCA CGCGCACCGTGTCATGCAAGCGCGGAGGATAGGCAAGTCCATGCTGGCA | 600 |
| DB | 541 | | |
| QY | 601 | CACAAGGTGATGTGGACTGGCGGAGCGGAACTCTTCCAAAGGCAGATTTGATTATCTC | 660 |
| DB | 601 | | |
| QY | 661 | TTCTACATCAACTGAGGAGATGAACCAAGAGTGCCACCGAATGCAAGACCTC | 720 |
| DB | 661 | | |
| QY | 721 | ATCTTCAGCTGCTGGCCTGAGCCAGCGCGCTCTCCAGGAGCTCATCCGAGTTCCGAG | 780 |
| DB | 721 | | |
| QY | 781 | CGCCTCCTTTTCATCATCGA CGGCTTCGATGAGCTCAAGCCTCTTTCCACGATCCTCAG | 840 |
| DB | 781 | | |
| QY | 841 | GGACCTGGTGCTCTGCTGGAGAGAAACCGGCCACCGAGCTGCTTTTAAcAGCTTA | 900 |
| DB | 841 | | |
| QY | 901 | ATTTCGGAAGAACTGCTCCCTGAGCTATCTTTGTCTATCAcCAcAGCGGCCACGGCTTGG | 960 |
| DB | 901 | | |
| QY | 961 | GAGAAGCTCCA CGGTCTGCTGGAGCAcCCcAGGcATGcGAGATcCTcGGcCTTCTcTGAG | 1020 |
| DB | 961 | | |
| QY | 1021 | GCAGAAGGAAGGATACTTCTACAAGTATTTCCAATGcAGAGcAGGcCGGcCAAGTC | 1080 |
| DB | 1021 | | |
| QY | 1081 | TTCAATTA CGTGAGGGAcAAcAGGcCTCTCTTCACCATGTCCTTCCTCGCTGGTGTC | 1140 |

| | | | |
|----|------|---|------|
| Db | 1081 | TTCAATTACGTAGGGACAAGAGCCTCTCTTCCACCATGTGCTTCTGTCGCCCTCGTGGTGC | 1140 |
| Qy | 1141 | TGGGTGGTGTGTACCTGCTCCAGCAGCAGCTCGAGGGTGGGGGCTGTGTGACACAGCG | 1200 |
| Db | 1141 | TGGGTGGTGTGTACCTGCTCCAGCAGCAGCTCGAGGGTGGGGGCTGTGTGACACAGCG | 1200 |
| Qy | 1201 | TCCAGGACCACCACTGCAGTGTACATGCTCTACTCTGCTGAGTCTGATGCAACCCAGCCG | 1260 |
| Db | 1201 | TCCAGGACCACCACTGCAGTGTACATGCTCTACTCTGCTGAGTCTGATGCAACCCAGCCG | 1260 |
| Qy | 1261 | GGGGCCCGCGCTTCAGCCCCCACCACAGAGAGGGTTGCTCTCTTGGGGCAGAT | 1320 |
| Db | 1261 | GGGGCCCGCGCTTCAGCCCCCACCACAGAGAGGGTTGCTCTCTTGGGGCAGAT | 1320 |
| Qy | 1321 | GGGCTCTGGAATCAGAAAAATCCTATTGAGGAGCAGGACCTCCGGAAGCAGCGGCTTAGAC | 1380 |
| Db | 1321 | GGGCTCTGGAATCAGAAAAATCCTATTGAGGAGCAGGACCTCCGGAAGCAGCGGCTTAGAC | 1380 |
| Qy | 1381 | GGGGAAGAGCTCTCGCTTCCTCAACATGAACATCTTCCAGNAGGACATCAACTGTGAG | 1440 |
| Db | 1381 | GGGGAAGAGCTCTCGCTTCCTCAACATGAACATCTTCCAGNAGGACATCAACTGTGAG | 1440 |
| Qy | 1441 | AGGTACTACAGCTTCATCCACTTGAGTTTCAGGAAATCTTTTGCAGCTATGTACTATATC | 1500 |
| Db | 1441 | AGGTACTACAGCTTCATCCACTTGAGTTTCAGGAAATCTTTTGCAGCTATGTACTATATC | 1500 |
| Qy | 1501 | CTGGAACGAGGGGAGGGCGGGGAGCCGACGACGAGGACGTCGACAGGCTGTTCACCGAG | 1560 |
| Db | 1501 | CTGGAACGAGGGGAGGGCGGGGAGCCGACGAGGACGTCGACAGGCTGTTCACCGAG | 1560 |
| Qy | 1561 | TACGGTGTCTGAAAGGAGCTTCTCTGGCACTCACCAGCGCTTCTCTGTTTGGACTCTCTG | 1620 |
| Db | 1561 | TACGGTGTCTGAAAGGAGCTTCTCTGGCACTCACCAGCGCTTCTCTGTTTGGACTCTCTG | 1620 |
| Qy | 1621 | AACGAGGAGACACGAGGCCACTTGGGAAGAGTCTCTGCTGGAAGTCTCGCGGCACATC | 1680 |
| Db | 1621 | AACGAGGAGACACGAGGCCACTTGGGAAGAGTCTCTGCTGGAAGTCTCGCGGCACATC | 1680 |
| Qy | 1681 | AAGATGGACCTGTTCAGTGGATCCAAAGCAAAGCTCAGAGGACGGCTCCACCCCTGCAG | 1740 |
| Db | 1681 | AAGATTGAACCTGTTCAGTGGATCCAAAGCAAAGCTCAGAGGACGGCTCCACCCCTGCAG | 1740 |
| Qy | 1741 | CAGGGCTCTTGGAGTCTTTCAGCTGTTGTAAGAGATCCAGGAGGAGAGTTCACCG | 1800 |
| Db | 1741 | CAGGGCTCTTGGAGTCTTTCAGCTGTTGTAAGAGATCCAGGAGGAGAGTTCACCG | 1800 |
| Qy | 1801 | CAGGGCTCTGAGCCACTTCAGTGCATCGTGTGAGCAATTCGCTCCAAAGATGGAGCAC | 1860 |
| Db | 1801 | CAGGGCTCTGAGCCACTTCAGTGCATCGTGTGAGCAATTCGCTCCAAAGATGGAGCAC | 1860 |
| Qy | 1861 | ATGTCTCTCTGTTCTGTGAAAGCTGACGAGGCGCCAGGTGTGTGATGGC | 1920 |
| Db | 1861 | ATGTCTCTCTGTTCTGTGAAAGCTGACGAGGCGCCAGGTGTGTGATGGC | 1920 |
| Qy | 1921 | GGCACTTACGCGCGGAGACCGCGCAGAGTCTCCGACGAGGCGCACACGCTG | 1980 |
| Db | 1921 | GGCACTTACGCGCGGAGACCGCGCAGAGTCTCCGACGAGGCGCACACGCTG | 1980 |
| Qy | 1981 | TTGGTGCAGCTCAGACACGAGAGGACCGTTCCTGTGAGCGCTTCACTGAACATCTGGCA | 2040 |
| Db | 1981 | TTGGTGCAGCTCAGACACGAGAGGACCGTTCCTGTGAGCGCTTCACTGAACATCTGGCA | 2040 |
| Qy | 2041 | GCGGCCCTGTGACCAATCCAACTGTATAGTGTCTGTATACCGAAATGCCCTGGGC | 2100 |
| Db | 2041 | GCGGCCCTGTGACCAATCCAACTGTATAGTGTCTGTATACCGAAATGCCCTGGGC | 2100 |
| Qy | 2101 | AGCCGGGGGTGAAAGTGTCTGTCAAGGACTTCAGACACCCCAACTGCAACTTCAGAAC | 2160 |
| Db | 2101 | AGCCGGGGGTGAAAGTGTCTGTCAAGGACTTCAGACACCCCAACTGCAACTTCAGAAC | 2160 |
| Qy | 2161 | CTGAGGCTGAAGGTGCCGATCTCCAGCTCAGCCTCGAGGACCTCTCTGCAAGCTCTC | 2220 |

```
Db 2161 CTGAGGCTGAAGAGTGCCGCGATCTCCAGCTCAGCCTGCGAGGACCTCTCTGCAAGCTCTC 2220
Qy 2221 ATAGCCAAATGAATTTGACAAGATGATCTCAGTGGCAACGCGGTGGATTCCTCCAGGC 2280
Db 2221 ATAGCCAAATGAATTTGACAAGATGATCTCAGTGGCAACGCGGTGGATTCCTCCAGGC 2280
Qy 2281 ATGATGCTGTTTTCGAGGGCTCGGCAATCCCAAGTGCAGGCTGCAGATGATTCAGTTG 2340
Db 2281 ATGATGCTGTTTTCGAGGGCTCGGCAATCCCAAGTGCAGGCTGCAGATGATTCAGTTG 2340
Qy 2341 AGGAAGTGTGAGCTGAGTGGGCTTGTGAGAGATGCTTCTGCTGCGGACCAAC 2400
Db 2341 AGGAAGTGTGAGCTGAGTGGGCTTGTGAGAGATGCTTCTGCTGCGGACCAAC 2400
Qy 2401 CCACATCTGTTGATTTGGACCTGACAGGAATCACTGGAAGATTTGGGCTCAGGTTA 2460
Db 2401 CCACATCTGTTGATTTGGACCTGACAGGAATCACTGGAAGATTTGGGCTCAGGTTA 2460
Qy 2461 CTATGCCAGGACTGAGGACCCAGTCTGAGACTAGCGACTTTGTGGCTGAAGATCTGC 2520
Db 2461 CTATGCCAGGACTGAGGACCCAGTCTGAGACTAGCGACTTTGTGGCTGAAGATCTGC 2520
Qy 2521 CGCCTCACTGCTGCTGTGAGAGTGGGCTCAACTCTCAGTGTGAACAGAGCTG 2580
Db 2521 CGCCTCACTGCTGCTGTGAGAGTGGGCTCAACTCTCAGTGTGAACAGAGCTG 2580
Qy 2581 AGAGAGTGGACCTGAGCTGAATGAGCTGGGACCTCGGGTGTCTGCTGTGTAG 2640
Db 2581 AGAGAGTGGACCTGAGCTGAATGAGCTGGGACCTCGGGTGTCTGCTGTGTAG 2640
Qy 2641 GGCTCAGGCAATCCAGTGAAGCTCGAGACCTGCGGTTGGGCACTCTGCCGCTGGC 2700
Db 2641 GGCTCAGGCAATCCAGTGAAGCTCGAGACCTCGGGTGTCTGCTGTGTAG 2700
Qy 2701 TCTGCCCTGTGAGGCTTCTTCTGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTG 2760
Db 2701 TCTGCCCTGTGAGGCTTCTTCTGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTG 2760
Qy 2761 GACTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGGTGTCTGCTGAGGCTGCA 2820
Db 2761 GACTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGGTGTCTGCTGAGGCTGCA 2820
Qy 2821 CATCCGCTGCACTCCAGAACTGTGGTGTGATGCTGTGGCTCAAGCAAGCT 2880
Db 2821 CATCCGCTGCACTCCAGAACTGTGGTGTGATGCTGTGGCTCAAGCAAGCT 2880
Qy 2881 TGTGAGATCTTTACTTCACTGCGGATCAACAGACCTTGACGACCTTTTACCTGAC 2940
Db 2881 TGTGAGATCTTTACTTCACTGCGGATCAACAGACCTTGACGACCTTTTACCTGAC 2940
Qy 2941 AACACGCTTAGGGGACACAGGTGTCCGACTGCTTTGCAAGCGGCTGAGCCATCTGGC 3000
Db 2941 AACACGCTTAGGGGACACAGGTGTCCGACTGCTTTGCAAGCGGCTGAGCCATCTGGC 3000
Qy 3001 TGAAACTCCAGTCTCTGTTATTTGGATGACCTGAAATAAATGACCCAGTAGG 3060
Db 3001 TGAAACTCCAGTCTCTGTTATTTGGATGACCTGAAATAAATGACCCAGTAGG 3060
Qy 3061 TTGGCAGGCTTCAGTAACAAACCTTATTTGACATTTGGCTCTGA 3108
Db 3061 TTGGCAGGCTTCAGTAACAAACCTTATTTGACATTTGGCTCTGA 3108
```

RESULT 2

```
US-10-407-866-23
; Sequence 23, Application US/10407866
; Publication No. US20040002593A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; AND METHODS OF USE
; FILE REFERENCE: 66654-10(LJ 5755)
```

```
; CURRENT APPLICATION NUMBER: US/10/407,866
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 60/370,538
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 3108
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3105)
US-10-407-866-23

Query Match 100.0%; Score 3108; DB 17; Length 3108;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTACGAACCGCAGCAGGACGCGCTCTGTGCTGTCCACTTCTGGAAGAACTC 60
Db 1 ATGCTACGAACCGCAGCAGGACGCGCTCTGTGCTGTCCACTTCTGGAAGAACTC 60
Qy 61 GAGGCTGTGAACCTGAAGAAATTTCAAGTTATACCTGCGGACCGGACAGAGCTGGAGAA 120
Db 61 GAGGCTGTGAACCTGAAGAAATTTCAAGTTATACCTGCGGACCGGACAGAGCTGGAGAA 120
Qy 121 GGCAAGATCCCTGGGGAAGCATGGAAAGCGGCTCCCTGGAAATGGCCAGCTGCTC 180
Db 121 GGCAAGATCCCTGGGGAAGCATGGAAAGCGGCTCCCTGGAAATGGCCAGCTGCTC 180
Qy 181 ATCACCACATTTCCGCGCAGAGGAGGCTGTGAGGTTGGCTCTCAGACACTTTGAGCGGATA 240
Db 181 ATCACCACATTTCCGCGCAGAGGAGGCTGTGAGGTTGGCTCTCAGACACTTTGAGCGGATA 240
Qy 241 AACAGGAGACCTGTGGGAGAGGACAGAGAGAGGACCTGTGAGGAGATCCCCAGGAA 300
Db 241 AACAGGAGACCTGTGGGAGAGGACAGAGAGAGGACCTGTGAGGAGATCCCCAGGAA 300
Qy 301 ACCTACAGGACATGTCCGACAGAAATTCGGCTCATGGAAGACCCCAATGCGCGCTA 360
Db 301 ACCTACAGGACATGTCCGACAGAAATTCGGCTCATGGAAGACCCCAATGCGCGCTA 360
Qy 361 GGGGAATGTCTAACTCAGCCACCGGTACACCGGCTCTCTGTGTGTAAGGAGCACTCA 420
Db 361 GGGGAATGTCTAACTCAGCCACCGGTACACCGGCTCTCTGTGTGTAAGGAGCACTCA 420
Qy 421 AACCCCATGAGGTCAGGACGAGCTTCTGGAACAGCCGCGGACACGCGAGGACCGTG 480
Db 421 AACCCCATGAGGTCAGGACGAGCTTCTGGAACAGCCGCGGACACGCGAGGACCGTG 480
Qy 481 GGACACGAGCTAGCCCATCAAGATAGACCTCTTTGAGCCAGACGAGGAGCGCCCC 540
Db 481 GGACACGAGCTAGCCCATCAAGATAGACCTCTTTGAGCCAGACGAGGAGCGCCCC 540
Qy 541 GAGCCACCGCGCACCGTGTGTCATGCAAGGCGCGGAGGATAGGCAAGTCCATCTGGCA 600
Db 541 GAGCCACCGCGCACCGTGTGTCATGCAAGGCGCGGAGGATAGGCAAGTCCATCTGGCA 600
Qy 601 CACAAGGTGATGTGGAATGCGGCGGAGAGCTTTCGAAGGAGCTTTCGAAGGAGATTTATCTC 660
Db 601 CACAAGGTGATGTGGAATGCGGCGGAGAGCTTTCGAAGGAGCTTTCGAAGGAGATTTATCTC 660
Qy 661 TTCTTACATCAACTGCGAGGAGATGAACAGAGTGCACGGAAATGACATGCAAGAGCTC 720
Db 661 TTCTTACATCAACTGCGAGGAGATGAACAGAGTGCACGGAAATGACATGCAAGAGCTC 720
Qy 721 ATCTTACAGCTGCTGGCTGAGCCGCGCTCTCCAGGAGCTCATCCGAGTTCCCGAG 780
Db 721 ATCTTACAGCTGCTGGCTGAGCCGCGCTCTCCAGGAGCTCATCCGAGTTCCCGAG 780
Qy 781 CGGCTCTTTTCAATCATGCAAGGCTTTCAGTGAAGCTTCTTTTCCAGGATCTCTTCCAG 840
Db 781 CGGCTCTTTTCAATCATGCAAGGCTTTCAGTGAAGCTTCTTTTCCAGGATCTCTTCCAG
```

| | | | |
|----|------|--|------|
| Db | 781 | CGCTCTCTTTTTCATCATCGACGGCTTCGATGAGCTCAAGCCTTCTTTTCCACGATCCTCMG | 840 |
| Qy | 841 | GGACCTCTGTGCTCTGCTGGAGAGAAACGGCCACGAGAGCTGCTTCTTAACAGCTTA | 900 |
| Db | 841 | GGACCTCTGTGCTCTGCTGGAGAGAAACGGCCACGAGAGCTGCTTCTTAACAGCTTA | 900 |
| Qy | 901 | ATTTCGGAAGAAGCTGCTCCCTGAGCTATCTTTTGTCTATCTACACAGGCCACGGCTTTG | 960 |
| Db | 901 | ATTTCGGAAGAAGCTGCTCCCTGAGCTATCTTTTGTCTATCTACACAGGCCACGGCTTTG | 960 |
| Qy | 961 | GAGAAGCTCCACCGTCTGCTGGAGCACCACGAGCATGTGGAGATCCTTGGGCTTCTCTGAG | 1020 |
| Db | 961 | GAGAAGCTCCACCGTCTGCTGGAGCACCACGAGCATGTGGAGATCCTTGGGCTTCTCTGAG | 1020 |
| Qy | 1021 | GCAGAAAGGAAGAACTTCTTCAAGTATTTTCACAATGCAGAGCAGGGCGGCCAAGTC | 1080 |
| Db | 1021 | GCAGAAAGGAAGAACTTCTTCAAGTATTTTCACAATGCAGAGCAGGGCGGCCAAGTC | 1080 |
| Qy | 1081 | TTCAATTAAGTGAAGGACAAAGAGCCTCTTCAACCATGTGCTTCCCTCCCTGTGTGC | 1140 |
| Db | 1081 | TTCAATTAAGTGAAGGACAAAGAGCCTCTTCAACCATGTGCTTCCCTCCCTGTGTGC | 1140 |
| Qy | 1141 | TGGGTGTGTGTACCTGCTCCAGCAGCAGCTGGAGGGTGGGGCTGTGTGAGACAGACG | 1200 |
| Db | 1141 | TGGGTGTGTGTACCTGCTCCAGCAGCAGCTGGAGGGTGGGGCTGTGTGAGACAGACG | 1200 |
| Qy | 1201 | TCCAGACCAACCTGAGTGTATCTTACCTGTGAGTCTGATGCCAACCCAGCCG | 1260 |
| Db | 1201 | TCCAGACCAACCTGAGTGTATCTTACCTGTGAGTCTGATGCCAACCCAGCCG | 1260 |
| Qy | 1261 | GGGGCCCGGCTCCAGCCCCCACCACACAGAGAGGGTGTGCTCTTGGGGCAGAT | 1320 |
| Db | 1261 | GGGGCCCGGCTCCAGCCCCCACCACACAGAGAGGGTGTGCTCTTGGGGCAGAT | 1320 |
| Qy | 1321 | GGGCTCTGGAATCAGAAAAATCTTATTTGAGAGAGAGACCTCCGGAGCAGGGCTAGAC | 1380 |
| Db | 1321 | GGGCTCTGGAATCAGAAAAATCTTATTTGAGAGAGAGACCTCCGGAGCAGGGCTAGAC | 1380 |
| Qy | 1381 | GGGAGAGAGCTCTGCTTCTTCAACATGAACATCTTCCAGAAAGACATCAACTGTGAG | 1440 |
| Db | 1381 | GGGAGAGAGCTCTGCTTCTTCAACATGAACATCTTCCAGAAAGACATCAACTGTGAG | 1440 |
| Qy | 1441 | AGGTACTACAGCTTCACTGAGTTTCCAGAGATTTCTTGGAGCTATGTAATATC | 1500 |
| Db | 1441 | AGGTACTACAGCTTCACTGAGTTTCCAGAGATTTCTTGGAGCTATGTAATATC | 1500 |
| Qy | 1501 | CTGACGAGGGGAGGGCGGGCAGCCAGACGAGAGCTGACAGAGCTGTGACCGAG | 1560 |
| Db | 1501 | CTGACGAGGGGAGGGCGGGCAGCCAGACGAGAGCTGACAGAGCTGTGACCGAG | 1560 |
| Qy | 1561 | TACGGCTTTTCTGAAAGAGCTTCTTGGCACTCACAGCGCTTCTGTTTGGACTCTG | 1620 |
| Db | 1561 | TACGGCTTTTCTGAAAGAGCTTCTTGGCACTCACAGCGCTTCTGTTTGGACTCTG | 1620 |
| Qy | 1621 | AACGAGGAGACGAGGACCTGAGAGAGTCTGCTGGAAGTCTCGCGGCACATC | 1680 |
| Db | 1621 | AACGAGGAGACGAGGACCTGAGAGAGTCTCTGCTGGAAGTCTCGCGGCACATC | 1680 |
| Qy | 1681 | AAGATGGAACCTGTGTCAGTGGATCCAAAGCAAAAGCTCAGAGCGACGGCTTCCACCTGAG | 1740 |
| Db | 1681 | AAGATGGAACCTGTGTCAGTGGATCCAAAGCAAAAGCTCAGAGCGACGGCTTCCACCTGAG | 1740 |
| Qy | 1741 | CAGGGCTCTTGGAGTCTTTCAGCTGTTGTAAGAGATCCAGATCCAGAGGAGAGTTTATCCAG | 1800 |
| Db | 1741 | CAGGGCTCTTGGAGTCTTTCAGCTGTTGTAAGAGATCCAGATCCAGAGGAGAGTTTATCCAG | 1800 |
| Qy | 1801 | CAGGGCTCTTGGAGTCTTTCAGCTGTTGTAAGAGATCCAGATCCAGAGGAGAGTTTATCCAG | 1860 |
| Db | 1801 | CAGGGCTCTTGGAGTCTTTCAGCTGTTGTAAGAGATCCAGATCCAGAGGAGAGTTTATCCAG | 1860 |
| Qy | 1861 | ATGGTCTCTGTTCTGTAAGGGCTGACAGGCGCCAGGTGCTGACCTTGTATGGC | 1920 |
| Db | 1861 | ATGGTCTCTGTTCTGTAAGGGCTGACAGGCGCCAGGTGCTGACCTTGTATGGC | 1920 |

Qy 3001 TGCAACTCCGAGTCTCTGTTATTTGGGATGGAAGTAAATGACCAAGTAGG 3060
Db 3001 TGCAACTCCGAGTCTCTGTTATTTGGGATGGAAGTAAATGACCAAGTAGG 3060
Qy 3061 TTGGCAGCGCTTCGAGTAACAAAACCTTATTTGGACATTTGCTCTGA 3108
Db 3061 TTGGCAGCGCTTCGAGTAACAAAACCTTATTTGGACATTTGCTCTGA 3108

RESULT 3

US-10-781-294-23
; Sequence 23, Application US/10781294
; Publication No. US20040142374A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Chu, Zhi-Liang
; APPLICANT: Pawlawski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Ariza, Maria Eugenia
; APPLICANT: Stehlik, Christian
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: P-LJ 4816
; CURRENT APPLICATION NUMBER: US/10/781,294
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US/09/965, 621
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/671, 760
; PRIOR FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 3108
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3105)
US-10-781-294-23

Query Match 100.0%; Score 3108; DB 18; Length 3108;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTACGAACCGCAGGAGGAGCGGCTCTGTGCGCTGTCCACTTACTTGGAGAACTC 60
Db 1 ATGCTACGAACCGCAGGAGGAGCGGCTCTGTGCGCTGTCCACTTACTTGGAGAACTC 60
Qy 61 GAGGCTGTGGAACCTGAAGAGTTTCAAGTTATACCTTGGGAGCCGCGACAGAGCTGGAGAA 120
Db 61 GAGGCTGTGGAACCTGAAGAGTTTCAAGTTATACCTTGGGAGCCGCGACAGAGCTGGAGAA 120
Qy 121 GCGAAGATCCCTGGGAGGAGATGAGAGGCGGCTCTGGAAGATGGCCAGCTGCTC 180
Db 121 GCGAAGATCCCTGGGAGGAGATGAGAGGCGGCTCTGGAAGATGGCCAGCTGCTC 180
Qy 181 ATCAACCACTTCGGGCGCAGAGAGGAGCTTGGAGTTGGCTCTCAGACCTTTGAGCGGATA 240
Db 181 ATCAACCACTTCGGGCGCAGAGAGGAGCTTGGAGTTGGCTCTCAGACCTTTGAGCGGATA 240
Qy 241 AACAGGAGGAGCTGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
Db 241 AACAGGAGGAGCTGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
Qy 301 ACCTACAGGAGGAGCTGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Db 301 ACCTACAGGAGGAGCTGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
Qy 361 GGGGAATGTGTCAACCTCAGCCAGCGGTACACCCGGCTCTGCTGGTGTGAGGAGCACTCA 420
Db 361 GGGGAATGTGTCAACCTCAGCCAGCGGTACACCCGGCTCTGCTGGTGTGAGGAGCACTCA 420

Qy 421 AACCCCATGACGAGTCCAGCAGCAGCTTCTGGACACAGCGCGGGACACGCGAGACCGTG 480
Db 421 AACCCCATGACGAGTCCAGCAGCAGCTTCTGGACACAGCGCGGGACACGCGAGACCGTG 480
Qy 481 GGACACAGGCTAGCCCCCATCAAGATAGACACCTCTTTGAGCCAGAGGAGGAGCCCC 540
Db 481 GGACACAGGCTAGCCCCCATCAAGATAGACACCTCTTTGAGCCAGAGGAGGAGCCCC 540
Qy 541 GAGCCACCGCGCACCGGTGTCATGCAAGGCGCGCAGGAGTAGGCAAGTCCATGCTGGCA 600
Db 541 GAGCCACCGCGCACCGGTGTCATGCAAGGCGCGCAGGAGTAGGCAAGTCCATGCTGGCA 600
Qy 601 CACAAGGTGATGTGGACTGGGCGGACGGGAAGCTTTTCCAAAGCAGATTTGATATCTC 660
Db 601 CACAAGGTGATGTGGACTGGGCGGACGGGAAGCTTTTCCAAAGCAGATTTGATATCTC 660
Qy 661 TTCTACATCAACTGCGAGGAGATGAACAGAGTGCCACGGAATGCGACATGCAAGACCTC 720
Db 661 TTCTACATCAACTGCGAGGAGATGAACAGAGTGCCACGGAATGCGACATGCAAGACCTC 720
Qy 721 ATCTTACAGTCTGGCTGAGCCCGCCTCTCCAGGAGCTCATCCGAGTTCCTCCGAG 780
Db 721 ATCTTACAGTCTGGCTGAGCCCGCCTCTCCAGGAGCTCATCCGAGTTCCTCCGAG 780
Qy 781 CGGCTCCTTTTTCATCATCGACGGCTTCGATGAGCTCAAGCTTCTTTCCACGATCTCAG 840
Db 781 CGGCTCCTTTTTCATCATCGACGGCTTCGATGAGCTCAAGCTTCTTTCCACGATCTCAG 840
Qy 841 GGACCTGTGCTCTGCTGGGAGGAGAAACGGCCACGGAGCTGCTTTTAAACAGCTTA 900
Db 841 GGACCTGTGCTCTGCTGGGAGGAGAAACGGCCACGGAGCTGCTTTTAAACAGCTTA 900
Qy 901 ATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACCAACGCGCCACCGCTTTG 960
Db 901 ATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACCAACGCGCCACCGCTTTG 960
Qy 961 GAGAAGCTCAACCGTCTGCTGGAGCAACCGAGGAGCTTGGAGATCTCTGGGCTTCTGAG 1020
Db 961 GAGAAGCTCAACCGTCTGCTGGAGCAACCGAGGAGCTTGGAGATCTCTGGGCTTCTGAG 1020
Qy 1021 GCAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Db 1021 GCAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Qy 1081 TTCAATTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db 1081 TTCAATTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Qy 1141 TGGGTGTGTGTACCTGCTCCAGCAGCAGCTGGAGGGTGGGGGCTGTTGAGACAGAG 1200
Db 1141 TGGGTGTGTGTACCTGCTCCAGCAGCAGCTGGAGGGTGGGGGCTGTTGAGACAGAG 1200
Qy 1201 TCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Db 1201 TCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Qy 1261 GGGGCCCCCGGCTCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db 1261 GGGGCCCCCGGCTCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Qy 1321 GGGCTCTGGAATCAGAAAATCTTATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
Db 1321 GGGCTCTGGAATCAGAAAATCTTATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
Qy 1381 GGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db 1381 GGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Qy 1441 AGGTACTACAGCTTCACTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
Db 1441 AGGTACTACAGCTTCACTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
Qy 1501 CTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560

| | | | | | | | |
|----|------|---|------|----|------|---|------|
| Qy | 610 | CAAGCAGATTGATTAATCTCTTCTTACATCAACTGCAGGAGATGAACACAGAGTGCCACG | 699 | Qy | 1720 | AGCGACGGCTCCACCTGCGACGAGGGCTCTCTTGAGTTCCTTTCAGCTGCTTGTACAGATC | 1779 |
| Db | 721 | CAAGCAGATTGATTAATCTCTTCTTACATCAACTGCAGGAGATGAACACAGAGTGCCACG | 780 | Db | 1801 | AGCGACGGCTCCACCTGCGACGAGGGCTCTCTTGAGTTCCTTTCAGCTGCTTGTACAGATC | 1860 |
| Qy | 700 | GAATGCAGCATGCAAGACCTCATCTTTCAGCTGCTGGCCCTGAGCCAGGCGCCCTCTCCAG | 759 | Qy | 1780 | CAGGAGGAGGTTTATCCAGCAGGGCCCTGAGCCACTTCCAGGTGATCGTGGTCAGCAAC | 1839 |
| Db | 781 | GAATGCAGCATGCAAGACCTCATCTTTCAGCTGCTGGCCCTGAGCCAGGCGCCCTCTCCAG | 840 | Db | 1861 | CAGGAGGAGGTTTATCCAGCAGGGCCCTGAGCCACTTCCAGGTGATCGTGGTCAGCAAC | 1920 |
| Qy | 760 | GAGCTCATCCGAGTTCGCGAGCGCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAG | 819 | Qy | 1840 | ATTGCCCTCCAGATGAGCAGCATGCTCTCTCTGTTCTGTCTGAAAGCGCTGCAGAGCGCC | 1899 |
| Db | 841 | GAGCTCATCCGAGTTCGCGAGCGCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAG | 900 | Db | 1921 | ATTGCCCTCCAGATGAGCAGCATGCTCTCTCTGTTCTGTCTGAAAGCGCTGCAGAGCGCC | 1980 |
| Qy | 820 | CCTTCTTTCCACGATCTCTCAGGACCCCTGGTGCCTCTGCTGGGAGGAGAAACGCGCCACG | 879 | Qy | 1900 | CAGGTGCTGCACTTTGATGCGCCACTTACAGCGCGACCGGGGAGAACCGCGCAGGTGC | 1959 |
| Db | 901 | CCTTCTTTCCACGATCTCTCAGGACCCCTGGTGCCTCTGCTGGGAGGAGAAACGCGCCACG | 960 | Db | 1981 | CAGGTGCTGCACTTTGATGCGCCACTTACAGCGCGACCGGGGAGAACCGCGCAGGTGC | 2040 |
| Qy | 880 | GAGCTGCTTCTTAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTCGTCATC | 939 | Qy | 1960 | TCCGACGAGGCGACAGCTGTTGGTGCAGCTCAGACACAGAGAGGACCGTTCCTGCTGGAC | 2019 |
| Db | 961 | GAGCTGCTTCTTAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTCGTCATC | 1020 | Db | 2041 | TCCGACGAGGCGCACAGCTGTTGGTGCAGCT---ACCAGAGAGGACCGTTCCTGCTGGAC | 2097 |
| Qy | 940 | ACCACAGCGCCACCGCTTTGGAGAGCTCCACCGTCTGCTGGAGACCCCGCAGGCATGTG | 999 | Qy | 2020 | GCCTACAGTGAACATCTGGCAGCGCCCTGTGCACCAATCCAAACCTGTATAGAGCTGTCT | 2079 |
| Db | 1021 | ACCACAGCGCCACCGCTTTGGAGAGCTCCACCGTCTGCTGGAGACCCCGCAGGCATGTG | 1080 | Db | 2098 | GCCTACAGTGAACATCTGGCAGCGCCCTGTGCACCAATCCAAACCTGTATAGAGCTGTCT | 2157 |
| Qy | 1000 | GAGATCCTGGGCTTCTCTGAGGCGAGAAAGGAGGAATCTTCTACAAGTATTTCCACAAT | 1059 | Qy | 2080 | CTGTACCGAAATGCTTGGGACGCGGGGGTGAAGCTGCTCTCAAGGACTCAGACAC | 2139 |
| Db | 1081 | GAGATCCTGGGCTTCTCTGAGGCGAGAAAGGAGGAATCTTCTACAAGTATTTCCACAAT | 1140 | Db | 2158 | CTGTACCGAAATGCTTGGGACGCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACAC | 2217 |
| Qy | 1060 | GCAGAGAGGGCGGCAAGTCTTCAATTAAGTGAAGGACAAAGAGCGCTCTCTCACCATG | 1119 | Qy | 2140 | CCCAACTGCAAACTTTCAGAACCTTGAGGCTGAGAGGTGCGCATCTCCAGCTCAGCCCTG | 2199 |
| Db | 1141 | GCAGAGAGGGCGGCAAGTCTTCAATTAAGTGAAGGACAAAGAGCGCTCTCTCACCATG | 1200 | Db | 2218 | CCCAACTGCAAACTTTCAGAACCTTGAGGCTGAGAGGTGCGCATCTCCAGCTCAGCCCTG | 2277 |
| Qy | 1120 | TGCTTCTGCTCCCTGCTGCTGGTGGTGTGTAACCTTCCAGCAGCAGCTGAGGGT | 1179 | Qy | 2200 | GAGGACCTCTCTGAGCTCTCATAGCCAAATAAGAAATTTGACAGATGAGTCTCAGTGGC | 2259 |
| Db | 1201 | TGCTTCTGCTCCCTGCTGCTGGTGGTGTGTAACCTTCCAGCAGCAGCTGAGGGT | 1260 | Db | 2278 | GAGGACCTCTCTGAGCTCTCATAGCCAAATAAGAAATTTGACAGATGAGTCTCAGTGGC | 2337 |
| Qy | 1180 | GGGGGGCTGTTGAGACAGACGCTCCAGGACCAACCTGAGTGTATCATGCTCTACCTGCTG | 1239 | Qy | 2260 | AACGGCTTGGAATCCAGGATGATGCTGCTTTCGAGGGCTGCGGCATCCCGAGTGC | 2319 |
| Db | 1261 | GGGGGGCTGTTGAGACAGACGCTCCAGGACCAACCTGAGTGTATCATGCTCTACCTGCTG | 1320 | Db | 2338 | AACGGCTTGGAATCCCGGATGATGCTGCTTTCGAGGGCTGCGGCATCCCGAATGC | 2397 |
| Qy | 1240 | AGTCTGATGCAACCGAGCGGGGCCCCGGCTCCAGCCCCCAGCCCAACAGAGAGG | 1299 | Qy | 2320 | AGGCTGAGATGATTCAGTTGAGAAAGTGTGAGTGGAGTCCGGGCTTGTTCAGAGATG | 2379 |
| Db | 1321 | AGTCTGATGCAACCGAGCGGGGCCCCGGCTCCAGCCCCCAGCCCAACAGAGAGG | 1380 | Db | 2398 | AGGCTGAGATGATTCAGTTGAGAAAGTGTGAGTGGAGTCCGGGCTTGTTCAGAGATG | 2457 |
| Qy | 1300 | TTGTGCTCCTTGGGCGAGATGGGCTCTGGAATCAGAAATCCTATTTGAGGACAGGAC | 1359 | Qy | 2380 | GCTTCTGCTGCTGGCACCAACCCACATCTGTTGAGTTGAGACCTGACAGGAAATGCACCTG | 2439 |
| Db | 1381 | TTGTGCTCCTTGGGCGAGATGGGCTCTGGAATCAGAAATCCTATTTGAGGACAGGAC | 1440 | Db | 2458 | GCTTCTGCTGCTGGCACCAACCCACATCTGTTGAGTTGAGACCTGACAGGAAATGCACCTG | 2517 |
| Qy | 1360 | CTCGGGAAGCAGCGCTAGACGGGGAAGACGCTCTGCTTCTTCTCAACATGAACATCTTC | 1419 | Qy | 2440 | GAGGATTTGGGCTGAGGTTTACTATGCCAGGGACTGAGGCAACCCAGTCTGCAGACTACGG | 2499 |
| Db | 1441 | CTCGGGAAGCAGCGCTAGACGGGGAAGACGCTCTGCTTCTTCTCAACATGAACATCTTC | 1500 | Db | 2518 | GAGGATTTGGGCTGAGGTTTACTATGCCAGGGACTGAGGCAACCCAGTCTGCAGACTACGG | 2577 |
| Qy | 1420 | CAGAGGACATCAACTGTGAGAGTATPACAGCTTATCCACTGAGTTTCAGGAATTC | 1479 | Qy | 2500 | ACTTTTGGCTGAAGATCTGCGGCTCATCTGCTGCTGCTGTGAGAGCTGGGCTTCAACT | 2559 |
| Db | 1501 | CAGAGGACATCAACTGTGAGAGTATPACAGCTTATCCACTTGAAGTTTCAGGAATTC | 1560 | Db | 2578 | ACTTTTGGCTGAGATCTGCGGCTCATCTGCTGCTGCTGTGAGAGCTGGGCTTCAACT | 2637 |
| Qy | 1480 | TTTGCAGCTATGACTATATCTCGACGAGGGGAGGGCGGGGAGGCGCCAGACAGGAC | 1539 | Qy | 2560 | CTCAGTCTGAACACAGACCTGAGAGAGCTGGAACCTGAGCTGGAATGAGCTGGGGAACCTC | 2619 |
| Db | 1561 | TTTGCAGCTATGACTATATCTCGACGAGGGGAGGGCGGGGAGGCGCCAGACAGGAC | 1620 | Db | 2638 | CTCAGTCTGAACACAGACCTGAGAGAGCTGGAACCTGAGCTGGAATGAGCTGGGGAACCTC | 2697 |
| Qy | 1540 | GTGACACGGCTGTTGACCGAGTACGCGTTTCTGAAAGGAGCTTCTTGGCACTCACGAGC | 1599 | Qy | 2620 | GGGCTGCTGCTGCTGTGAGGGGCTCAGGCATCCCACGTGCAAGCTCCAGACCTCGCG | 2679 |
| Db | 1621 | GTGACACGGCTGTTGACCGAGTACGCGTTTCTGAAAGGAGCTTCTTGGCACTCACGAGC | 1680 | Db | 2698 | GGGCTGCTGCTGCTGTGAGGGGCTCAGGCATCCCACGTGCAAGCTCCAGACCTCGCG | 2757 |
| Qy | 1600 | CGCTTCTGTTTGAATCCTGAAACGAGAGACAGGAGCCACCTGGAGAGAGTCTCTGC | 1659 | Qy | 2680 | TTGGGCACTCTGCGGCTGGGCTCTGCGGCTGTGAGGGTCTTCTGTGTGTGCTCAGGCC | 2739 |
| Db | 1681 | CGCTTCTGTTTGAATCCTGAAACGAGAGACAGGAGCCACCTGGAGAGAGTCTCTGC | 1740 | Db | 2758 | TTGGGCACTCTGCGGCTGGGCTCTGCGGCTGTGAGGGTCTTCTGTGTGTGCTCAGGCC | 2817 |
| Qy | 1660 | TGGAGGTCTCGCGCACATCAAGATGACCTGTTGAGTGGATCCAAAGCAAGAGCTCAG | 1719 | Qy | 2740 | AACCAACAACCTCCGGGAGCTGGAATTCAGTTCCTGAGGAGTCTGGGCTGTGG | 2799 |
| Db | 1741 | TGGAGGTCTCGCGCACATCAAGATGACCTGTTGAGTGGATCCAAAGCAAGAGCTCAG | 1800 | Db | 2818 | AACCAACAACCTCCGGGAGCTGGAATTCAGTTCCTGAGTTCCTGAGGAGTCTGGGCTGTGG | 2877 |
| | | | | Qy | 2800 | TTGCTGCTGAGGGGCTGCAACATCCCGCTGACAGACTCCAGAAACTGTGTGCTGGATAGC | 2859 |

| | | |
|------|--|------|
| 181 | CCACCCAAACAGAGAGGGTTGTCTCTCTTGGCGGAGATGGGCTCTCGAATCAGAAATC | 240 |
| 1342 | CTATTTGAGGAGCAGGACCTCCGGAAGACACGGCCTTAGACGGGGAAGACGTCCTTGCCTTC | 1401 |
| 241 | CTATTTGAGGAGCAGGACCTCCGGAAGACACGGCCTTAGACGGGGAAGACGTCCTTGCCTTC | 300 |
| 1402 | CTCAAATGAAACATCTTCAGAAAGACATCAACTGTGAGAGGTACTACAGCTTCATCCAC | 1461 |
| 301 | CTCAAATGAAACATCTTCAGAAAGACATCAACTGTGAGAGGTACTACAGCTTCATCCAC | 360 |
| 1462 | TTGAGTTTCCAGGAATTCCTTGCACTATGACTATATCTCTGAGACGAGGGGAGGGCGGG | 1521 |
| 361 | TTGAGTTTCCAGGAATTCCTTGCACTATGACTATATCTCTGAGACGAGGGGAGGGCGGG | 420 |
| 1522 | GCAGGCCACACAGGACGTCGACCGAGCTGTTGACCGAGTACGGTTCCTGAAAGAGAC | 1581 |
| 421 | GCAGGCCACACAGGACGTCGACCGAGCTGTTGACCGAGTACGGTTCCTGAAAGAGAC | 480 |
| 1582 | TTCTGTGCACTACACGACGGCTTCCTGTTTGGACTCTTGAAACGAGGAGACACGAGGCAC | 1641 |
| 481 | TTCTGTGCACTACACGACGGCTTCCTGTTTGGACTCTTGAAACGAGGAGACACGAGGCAC | 540 |
| 1642 | CTGAGAAAGAGTCTCTGCTGGAAAGTCTCCGCCACATCAAGATGGACCTGTTGCAGTGG | 1701 |
| 541 | CTGAGAAAGAGTCTCTGCTGGAAAGTCTCCGCCACATCAAGATGGACCTGTTGCAGTGG | 600 |
| 1702 | ATCCAAAGAAAGCTCAGAGCAGCGCTCCACCTCTGACGAGGGCTCCTTGGAGTTCTTC | 1761 |
| 601 | ATCCAAAGAAAGCTCAGAGCAGCGCTCCACCTCTGACGAGGGCTCCTTGGAGTTCTTC | 660 |
| 1762 | AGCTGCTTTACAGATCCAGAGGAGGAGTTCATCCAGCAGGCCCTGAGGCCATCTCCAG | 1821 |
| 661 | AGCTGCTTTACAGATCCAGAGGAGGAGTTCATCCAGCAGGCCCTGAGGCCATCTCCAG | 720 |
| 1822 | GTGATCTGTGTCAGCAACATTTGCCTTCAGATGGAGCAATGGTCTCCTCGTCTGTCTG | 1881 |
| 721 | GTGATCTGTGTCAGCAACATTTGCCTTCAGATGGAGCAATGGTCTCCTCGTCTGTCTG | 780 |
| 1882 | AAGCGCTGCAGAGCGCCAGGTCTGTGCACTGTATGGCGCCACCTACAGCGCGGACGGG | 1941 |
| 781 | AAGCGCTGCAGAGCGCCAGGTCTGTGCACTGTATGGCGCCACCTACAGCGCGGACGGG | 840 |
| 1942 | GAAGACCGCGAGGTGCTCCGAGGAGCGCACACCGCTGTTGGTGCAGCTCAGACACGAG | 2001 |
| 841 | GAAGACCGCGAGGTGCTCCGAGGAGCGCACACCGCTGTTGGTGCAGCTCAGACACGAG | 900 |
| 2002 | AGGACCGTCTGCTGAGCGCTACAGTGAACATCTGGCAGCGCCCTGTGCAACCAATCCA | 2061 |
| 901 | AGGACCGTCTGCTGAGCGCTACAGTGAACATCTGGCAGCGCCCTGTGCAACCAATCCA | 960 |
| 2062 | AACCTGTATAGACTGTCTCTGTAACGAAATGCCCTTGGGACGCCGGGGGTGAAGCTGCTC | 2121 |
| 961 | AACCTGTATAGACTGTCTCTGTAACGAAATGCCCTTGGGACGCCGGGGGTGAAGCTGCTC | 1020 |
| 2122 | TGTCAAGACTCAGACACCCCAACTGCAAACTTCAGAACTCTGAGGAGGTGAGAGGTC | 2178 |
| 1021 | TGTCAAGACTCAGACACCCCAACTGCAAACTTCAGAACTCTGAGGAGGTGAGAGGTC | 1080 |
| 2179 | CGCATCTCAGCTCAGCGCTCGGAGGACCTCTCTGAGCTCTCATAGCCAATAGAAATTTG | 2238 |
| 1081 | CGCATCTCAGCTCAGCGCTCGGAGGACCTCTCTGAGCTCTCATAGCCAATAGAAATTTG | 1140 |
| 2239 | ACAAGATGGATCTCAGTGGCAACGGCGTTGGATTCACGAGCATGATGCTTTGCGAG | 2298 |
| 1141 | ACAAGATGGATCTCAGTGGCAACGGCGTTGGATTCACGAGCATGATGCTTTGCGAG | 1200 |
| 2299 | GGCTTCGGCATCCCCAGTGCAGGCTGCAGATGATTCAGTTGAGGAAGTGTCACTGGAG | 2358 |
| 1201 | GGCTTCGGCATCCCCAGTGCAGGCTGCAGATGATTCAGTTGAGGAAGTGTCACTGGAG | 1260 |
| 2359 | TCCGGGGCTTGTACAGAGATGCGCTTCTGTGCTCGGCACCAACCAATCTGTTGAGTTG | 2418 |
| 1261 | TCCGGGGCTTGTACAGAGATGCGCTTCTGTGCTCGGCACCAACCAATCTGTTGAGTTG | 1320 |

RESULT 9

```

1  RESUME:
2  US-10-467-397-19
3  ;
4  ; Sequence 19, Application US/10467397
5  ; Publication No. US20040137448A1
6  ;
7  ; GENERAL INFORMATION:
8  ;
9  ; APPLICANT: THORNTON, Michael; HAFALIA, Ap
10 ;
11 ; APPLICANT: LU, Dyung Aina M.; ARUVIZU, Cha
12 ;
13 ; APPLICANT: SWARNAKAR, Anita; LU, Yan;
14 ;
15 ; APPLICANT: WARREN, Bridget A.; BAUGHN, Me
16 ;
17 ; APPLICANT: TANG, Y. Tom; LEE, Ernestine J
18 ;
19 ; APPLICANT: YAO, Monique G.; RAKKUMAR, Jay
20 ;
21 ; APPLICANT: KHAN, Farrah A.; GANDHI, Ameet
22 ;
23 ; APPLICANT: DING, Li; YUB, Henry;
24 ;
25 ; APPLICANT: GIETZEN, Kimberly J.; CHAWLA,
26 ;
27 ; APPLICANT: THANGAVELU, Kavitha; ELLIOTT,
28 ;
29 ; APPLICANT: MARQUIS, Joseph P.
30 ;
31 ; TITLE OF INVENTION: NUCLEIC ACID-ASSOCIAT
32 ;
33 ; FILE REFERENCE: PI-0368 USN
34 ;
35 ; CURRENT APPLICATION NUMBER: US/10/467,397
36 ;
37 ; CURRENT FILING DATE: 2003-08-06
38 ;
39 ; PRIOR APPLICATION NUMBER: PCT/US02/03844
40 ;
41 ; PRIOR FILING DATE: 2002-02-07
42 ;
43 ; PRIOR APPLICATION NUMBER: US 60/268,118
44 ;
45 ; PRIOR FILING DATE: 2001-02-09
46 ;
47 ; PRIOR APPLICATION NUMBER: US 60/270,963
48 ;
49 ; PRIOR FILING DATE: 2001-02-21
50 ;
51 ; PRIOR APPLICATION NUMBER: US 60/270,858
52 ;
53 ; PRIOR FILING DATE: 2001-02-22
54 ;
55 ; PRIOR APPLICATION NUMBER: US 60/271,194
56 ;
57 ; PRIOR FILING DATE: 2001-02-23
58 ;
59 ; PRIOR APPLICATION NUMBER: US 60/274,071
60 ;
61 ; PRIOR FILING DATE: 2001-03-07
62 ;
63 ; PRIOR APPLICATION NUMBER: US 60/283,496
64 ;
65 ; PRIOR FILING DATE: 2001-04-12
66 ;
67 ; PRIOR APPLICATION NUMBER: US 60/344,650
68 ;
69 ; PRIOR FILING DATE: 2001-11-09
70 ;
71 ; NUMBER OF SEQ ID NOS: 20
72 ;
73 ; SOFTWARE: PERL Program
74 ;
75 ; SEQ ID NO 19

```

LENGTH: 2158
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 429930CB1
US-10-467-397-19

Query Match 54.1%; Score 1681.8; DB 18; Length 2158;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1696; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

Qy 1408 ATGAACATCTCCAGAGGACATCAACTGTGAGAGTACTACAGCTTCACTCAGTTGAGT 1467
Db 1 ATGAACATCTCCAGAGGACATCAACTGTGAGAGTACTACAGCTTCACTCAGTTGAGT 60
Qy 1468 TTCCAGGAATCTTTGGAGCTATGCTATATCTGACGAGGGGGGGGGGGGGGGGGGG 1527
Db 61 TTCCAGGAATCTTTGGAGCTATGCTATATCTGACGAGGGGGGGGGGGGGGGGGGG 120
Qy 1528 CCAGACAGGAGCTGACAGGCTGTTGACCGAGTACGCGTTTCTGAAAGGAGCTTCTCTG 1587
Db 121 CCAGACAGGAGCTGACAGGCTGTTGACCGAGTACGCGTTTCTGAAAGGAGCTTCTCTG 180
Qy 1588 GCATCTACAGCGGCTTCTGTTGGACTCTCGAAACGAGGAGACAGAGGACCACTGGAG 1647
Db 181 GCATCTACAGCGGCTTCTGTTGGACTCTCGAAACGAGGAGACAGAGGACCACTGGAG 240
Qy 1648 AAGAGTCTCTGCTGGAGGCTCTCCGACATCAAGATGAGCTGTTGCAAGTGAATCCAA 1707
Db 241 AAGAGTCTCTGCTGGAGGCTCTCCGACATCAAGATGAGCTGTTGCAAGTGAATCCAA 300
Qy 1708 AGCAAGCTCAGAGCGAGGCTCCACCTGACGAGGCTCTTGGAGTCTTCCAGCTGC 1767
Db 301 AGCAAGCTCAGAGCGAGGCTCCACCTGACGAGGCTCTTGGAGTCTTCCAGCTGC 360
Qy 1768 TTGTACAGATCCAGGAGGAGGATTTATCCAGCAGGCGCTTGACCACTTCCAGGTGATC 1827
Db 361 TTGTACAGATCCAGGAGGAGGATTTATCCAGCAGGCGCTTGACCACTTCCAGGTGATC 420
Qy 1828 GTGCTCAGACATTTGCTCCAGATGGAGCAGATGCTCTGCTGCTGCTGCTGCTGCTGCTG 1887
Db 421 GTGCTCAGACATTTGCTCCAGATGGAGCAGATGCTCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 1888 TGCAGGAGCGCCAGGCTGCTCACTTGTATGGCGCCACCTACAGCGGGGAGCGGGAAGAC 1947
Db 481 TGCAGGAGCGCCAGGCTGCTCACTTGTATGGCGCCACCTACAGCGGGGAGCGGGAAGAC 540
Qy 1948 CGCGGAGGTGCTCCGACAGGAGCGACACGCTGTTGTTGCTGACCTCAGACAGAGGAGC 2007
Db 541 CGCGGAGGTGCTCCGACAGGAGCGACACGCTGTTGTTGCTGACCT-...ACCAGAGAGGAC 597
Qy 2008 GTTCTGCTGGAGGCTACAGTGAACATCTGACGAGGCGCTGTCACCAATCCAAACCTG 2067
Db 598 GTTCTGCTGGAGGCTACAGTGAACATCTGACGAGGCGCTGTCACCAATCCAAACCTG 657
Qy 2068 ATAGAGTGTCTGTATCCGAAATGCCCTGGGAGCGGGGGTGAAGCTGCTCTGTCAA 2127
Db 658 ATAGAGTGTCTGTATCCGAAATGCCCTGGGAGCGGGGGTGAAGCTGCTCTGTCAA 717
Qy 2128 GGACTCAGACACCCCAACTGCAAACTTCAGAACCTGAGGCTGAAGAGGTGCCGATCTCC 2187
Db 718 GGACTCAGACACCCCAACTGCAAACTTCAGAACCTGAGGCTGAAGAGGTGCCGATCTCC 777
Qy 2188 AGCTCAGCTGGAGGAGCTCTGACGCTCTCATAGCAATAGAATTGACAAGATG 2247
Db 778 AGCTCAGCTGGAGGAGCTCTGACGCTCTCATAGCAATAGAATTGACAAGATG 837
Qy 2248 GATCTCAGTGGCAACGGGCTTGGATTTCCAGGATGATGCTGCTTGGAGGGCTCGGG 2307
Db 838 GATCTCAGTGGCAACGGGCTTGGATTTCCAGGATGATGCTGCTTGGAGGGCTCGGG 897
Qy 2308 CATCCCCAGTGCAGGCTSCAGATGATTCAGTTGAGGAAGTGTGACGTGGAGTCCGGGCT 2367

Db 898 CATCCCCAATGCGAGGCTGCAGATGATTCAGTTGAGGAAGTGTGACGTGGAGTCCGGGCT 957
Qy 2368 TGTCCAGGAGATGCTTCTGTGCTCGGACCAACCCACATCTGTTGAGTTGAGCTGAC 2427
Db 958 TGTCCAGGAGATGCTTCTGTGCTTGGACCAACCCACATCTGTTGAGTTGAGCTGAC 1017
Qy 2428 GGAATATGCTGAGGAGATTTGGGCTGAGGTTACTATGCCAGGAGCTGAGGACCCAGT 2487
Db 1018 GGAATATGCTGAGGAGATTTGGGCTGAGGTTACTATGCCAGGAGCTGAGGACCCAGT 1077
Qy 2488 TGCAGACTACGAGCTTGTGCTGAGACTGCGGCTCACTGCTGCTGCTGCTGCTGCTGCTG 2547
Db 1078 TGCAGACTACGAGCTTGTGCTGAGAGTCTGCGGCTCACTGCTGCTGCTGCTGCTGCTG 1137
Qy 2548 CTGCGCTCAACTCTCAGTGTGAACAGAGCTGAGAGAGCTGGAACCTGAGCTGGAATGAG 2607
Db 1138 CTGCGCTCAACTCTCAGTGTGAACAGAGCTGAGAGAGCTGGAACCTGAGCTGGAATGAG 1197
Qy 2608 CTGCGGAGCTCTGCGGCTGCTGCTGCTGAGAGGCTTCAAGGATCCAGTGAAGCTC 2667
Db 1198 CTGCGGAGCTCTGCGGCTGCTGCTGCTGAGGAGCTTCAAGGATCCAGGATCCAGGATC 1257
Qy 2668 CAGACCTCTGCGGCTTGGGCTCTGCGGCTGCGGCTCTGCGGCTGCTGAGGCTCTTCTG 2727
Db 1258 CAGACCTCTGCGGCTTGGGCTCTGCGGCTGCGGCTCTGCGGCTCTGAGGCTCTTCTG 1317
Qy 2728 GTGCTCAGGCGCCAAACCAACCTCCGGGAGCTGAGCTTGAAGTTCAACGAGCTGGAGAC 2787
Db 1318 GTGCTCAGGCGCCAAACCAACCTCCGGGAGCTGAGCTTGAAGTTCAACGAGCTGGAGAC 1377
Qy 2788 TGGGCGCTGTGCTGCTGCTGAGGCTGCAACATCCCGCTGACAGCTCCAGAACTG 2847
Db 1378 TGGGCGCTGTGCTGCTGCTGAGGCTGCAACATCCCGCTGACAGCTCCAGAACTG 1437
Qy 2848 TGGCTGATGAGCTGTGCGCTCAGGCGGCTTGTGAGATCTTTACTTACCTCAGCTGGG 2907
Db 1438 TGGCTGATGAGCTGTGCGCTCAGGCGGCTTGTGAGATCTTTACTTACCTCAGCTGGG 1497
Qy 2908 ATCAACAGAGCTTGTGCGGCTTGTGCTGCAACCAACCGCTTGTGAGGAGCAGAGGTG 2967
Db 1498 ATCAACAGAGCTTGTGCGGCTTGTGCTGCAACCAACCGCTTGTGAGGAGCAGAGGTG 1557
Qy 2968 CGACTGCTTTGCAAGCGGCTGAGCCATCTGCTGCTGCAAACTCCAGTCTCTGTTATTT 3027
Db 1558 CGACTGCTTTGCAAGCGGCTGAGCCATCTGCTGCTGCAAACTCCAGTCTCTGTTATTT 1617
Qy 3028 GGGATGAGCTGAATAAATGACCCAGTGTGCGGCTTGTGAGGCTTGTGAGGCTTGTGAG 3087
Db 1618 GGGATGAGCTGAATAAATGACCCAGTGTGCGGCTTGTGAGGCTTGTGAGGCTTGTGAG 1677
Qy 3088 TATTTGACATTTGCTGCTGA 3108
Db 1678 TATTTGACATTTGCTGCTGA 1698

RESULT 10

US-10-029-386-24660/c
; Sequence 24660, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Rank, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: EXPRISION ANALYSIS TWO
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24660
; LENGTH: 1704

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008753.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.37
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EST HUMAN HIT: AW673661.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P33076, EVALUATE 4.00e-13
; OTHER INFORMATION: NT HIT: g114731965, EVALUATE 1.00e-07
; US-10-029-386-24660

Query Match 54.1%; Score 1680; DB 16; Length 1704;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1702; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 290 ATCCCAAGGAACCTACAGGACTATGTCGCGAGGAATTCGCGCTCATGGAAGACGCA 349
DB 1704 ATCCCAAGGAACCTACAGGACTATGTCGCGAGGAATTCGCGCTCATGGAAGACGCA 1645
QY 350 ATGCGCGCTAGGGAAATGTGTCAACCTCAGCCACCGGTACACCCGGCTCCTGCTGTGA 409
DB 1644 ATGCGCGCTAGGGAAATGTGTCAACCTCAGCCACCGGTACACCCGGCTCCTGCTGTGA 1585
QY 410 AGGAGCACTCAAAACCCATGAGTCCAGACGAGCTTCTGGAACAGGCGGGGACAG 469
DB 1584 AGGAGCACTCAAAACCCATGAGTCCAGACGAGCTTCTGGAACAGGCGGGGACAG 1525
QY 470 CGAGGACCGTGGGACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGACG 529
DB 1524 CGAGGACCGTGGGACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGACG 1465
QY 530 AGGAGCGCCCGAGCCACCGGCAACCGTGTCTATGCAAGGCGCGGAGGATAGGCAAGT 589
DB 1464 AGGAGCGCCCGAGCCACCGGCAACCGTGTCTATGCAAGGCGCGGAGGATAGGCAAGT 1405
QY 590 CCATGCTGGCACACAGGTGATGCTGAGTGGCGGAGCGGAAGCTCTTCAAGGAGAT 649
DB 1404 CCATGCTGGCACACAGGTGATGCTGAGTGGCGGAGCGGAAGCTCTTCAAGGAGAT 1345
QY 650 TTGATTATCTCTTACATCAACTGCAGGAGATGAACAGAGTGCCACCGAAATGCAACA 709
DB 1344 TTGATTATCTCTTACATCAACTGCAGGAGATGAACAGAGTGCCACCGAAATGCAACA 1285
QY 710 TCCAAAGACTATCTTACGTGCTGCGCTGAGCCAGCGCGCTCTCCAGAGCTCATCC 769
DB 1284 TCCAAAGACTATCTTACGTGCTGCGCTGAGCCAGCGCGCTCTCCAGAGCTCATCC 1225
QY 770 GAGTCCCGAGCGCTCTCTTTTATCATCGAGGCTTCGATGAGCTCAAGCTCTTTTCC 829
DB 1224 GAGTCCCGAGCGCTCTCTTTTATCATCGAGGCTTCGATGAGCTCAAGCTCTTTTCC 1165
QY 830 ACGATCCTCAGGGAACCTGCTGCTCTGCGGAGGAAACCGGCCACCGAGCTGCTTC 889
DB 1164 ACGATCCTCAGGGAACCTGCTGCTCTGCGGAGGAAACCGGCCACCGAGCTGCTTC 1105
QY 890 TTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCAACACGCG 949
DB 1104 TTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCAACACGCG 1045
QY 950 CCACGCTTTGGAGAGCTCCACCGCTCTGCTGAGCACCCACCGCATGTGGAGATCTCTGG 1009
DB 1044 CCACGCTTTGGAGAGCTCCACCGCTCTGCTGAGCACCCACCGCATGTGGAGATCTCTGG 985
QY 1010 GCTTCTCTGAGGAGGAAGGAATACTTCTACAGTATTTCCACATGAGGAGGAGG 1069
DB 984 GCTTCTCTGAGGAGGAAGGAATACTTCTACAGTATTTCCACATGAGGAGGAGG 925
QY 1070 CGGGCAAGCTCTCAATTAGTGAGGAGCAACGAGCTCTCTTACCATGCTCTCTGCTCC 1129
DB 924 CGGGCAAGCTCTCAATTAGTGAGGAGCAACGAGCTCTCTTACCATGCTCTCTGCTCC 865
QY 1130 CCCTGTGTGCTGGGTGCTGTGTACCTGCTCCAGCAGCTGAGGAGTGGGGGCTGT 1189

DB 864 CCCTGTGTGCTGGGTGCTGTGTACCTGCTCCAGCAGCAGCTGGAGGGTGGGGGCTGT 805
QY 1190 TGAGACAGACGTCCAGGACCAACCTGAGTGTACATGCTCTACCTGCTGAGTGTATGC 1249
DB 804 TGAGACAGACGTCCAGGACCAACCTGAGTGTACATGCTCTACCTGCTGAGTGTATGC 745
QY 1250 AACCACAGCGGGGGCGCGCTCCAGCCCCCACCACCAACAGAGAGGGTGTGCTCCT 1309
DB 744 AACCACAGCGGGGGCGCGCTCCAGCCCCCACCACCAACAGAGAGGGTGTGCTCCT 685
QY 1310 TGGGGCAGATGGGCTCTGGAATCAGAAATCTTATTGAGGAGCAGGACCTCCGGAAGC 1369
DB 684 TGGGGCAGATGGGCTCTGGAATCAGAAATCTTATTGAGGAGCAGGACCTCCGGAAGC 625
QY 1370 ACGGCTTAGACGGGGAAGACGTCTCTGCTTCTCAACATGAACATCTTCCAGAAGACA 1429
DB 624 ACGGCTTAGACGGGGAAGACGTCTCTGCTTCTCAACATGAACATCTTCCAGAAGACA 565
QY 1430 TCAACTGTGAGAGGTACTACAGCTTCACTCTTGAATTTCCAGGAATTTCTTCAGCTA 1489
DB 564 TCAACTGTGAGAGGTACTACAGCTTCACTCTTGAATTTCCAGGAATTTCTTCAGCTA 505
QY 1490 TGTACTATATCTTGAACGAGGGGAGGCGGGGAGGCGGAGACGAGACGAGACGAGG 1549
DB 504 TGTACTATATCTTGAACGAGGGGAGGCGGGGAGGCGGAGACGAGACGAGACGAGG 445
QY 1550 TGTGACCGAGTACCGCTTTCTGAAAGGAGCTTCTGCACTCACCAGCGCTTCTCTGT 1609
DB 444 TGTGACCGAGTACCGCTTTCTGAAAGGAGCTTCTGCACTCACCAGCGCTTCTCTGT 385
QY 1610 TTGAGTCTCTGAAACGAGGAGACGAGGCGGAGGCGGAGGAGGAGGAGGAGGAGG 1669
DB 384 TTGAGTCTCTGAAACGAGGAGACGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 325
QY 1670 CCGCGCATCAAGATGAGACCTTTGAGTGTGATGATCCAAAGCAAGCTCAGAGCGAGCT 1729
DB 324 CCGCGCATCAAGATGAGACCTTTGAGTGTGATGATCCAAAGCAAGCTCAGAGCGAGCT 265
QY 1730 CCACCTGCGAGCAGGCTCTTTGAGTGTGATGATCCAAAGCAAGCTCAGAGGAGGAGG 1789
DB 264 CCACCTGCGAGCAGGCTCTTTGAGTGTGATGATCCAAAGCAAGCTCAGAGGAGGAGG 205
QY 1790 AGTTTATCAGCAGGCGCTTGAGCCACTTCCAGGTATCGTGTGAGCAACATTTGCTCCA 1849
DB 204 AGTTTATCAGCAGGCGCTTGAGCCACTTCCAGGTATCGTGTGAGCAACATTTGCTCCA 145
QY 1850 AGATGAGCACATGCTCTCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 1908
DB 144 AGATGAGCACATGCTCTCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 85
QY 1909 CACTTGTGTA-TGGCGCCACCTTACAGCGGAGCGGGGAGAACCGCGGAGGTGTCCGAGG 1967
DB 84 CACTTGTGTA-TGGCGCCACCTTACAGCGGAGCGGGGAGAACCGCGGAGGTGTCCGAGG 25
QY 1968 AGCGCACACGCTTTGCTGTCAGCT 1991
DB 24 AGCGCACACGCTTTGCTGTCAGCT 1

RESULT 11
US-10-028-374-1
; Sequence 1, Application US/10028374
; Publication No. US20030143706A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; FILE REFERENCE: D0067NP
; CURRENT APPLICATION NUMBER: US/10/028,374
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1

; LENGTH: 4931

; TYPE: DNA

; ORGANISM: homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1320)..(2666)

US-10-028-374-1

Query Match 52.0%; Score 1616; DB 15; Length 4931;
Best Local Similarity 81.0%; Pred. No. 0;
Matches 2160; Conservative 0; Mismatches 5; Indels 501; Gaps 4;

| | | | |
|----|------|--|------|
| Qy | 289 | GATCCCCAGGAAACCTCAGGAGCTATGTCCGCGAGAAATTCGGCTCATGGAAAGCCG | 348 |
| Db | 1021 | GATCCCCAGGAAACCTCAGGAGCTATGTCCGCGAGAAATTCGGCTCATGGAAAGCCG | 1080 |
| Qy | 349 | AATGGCGCCTTAGGGAAATGTCAACCTCAGCCACCGGTACACCGGCTCCTGTGGTG | 408 |
| Db | 1081 | AATGGCGCCTTAGGGAAATGTCAACCTCAGCCACCGGTACACCGGCTCCTGTGGTG | 1140 |
| Qy | 409 | AAGGAGCACTCAAAACCCCATCAGCTCCAGCAGCAGCTTCTGGACACAGCGCGGGACAC | 468 |
| Db | 1141 | AAGGAGCACTCAAAACCCCATCAGCTCCAGCAGCAGCTTCTGGACACAGCGCGGGACAC | 1200 |
| Qy | 469 | GCGAGGACCGTGGGACACAGGCTAGCCCCATCAAGATAGAGACCTCTTTGAGCCAGAC | 528 |
| Db | 1201 | GCGAGGACCGTGGGACACAGGCTAGCCCCATCAAGATAGAGACCTCTTTGAGCCAGAC | 1260 |
| Qy | 529 | GAGGAGCGCCCGAGCCACCGCGCACCGTGTCTATGCAAGCGCGGCGGAGGATAGGCAAG | 588 |
| Db | 1261 | GAGGAGCGCCCGAGCCACCGCGCACCGTGTCTATGCAAGCGCGGCGGAGGAT | 1313 |
| Qy | 589 | TCCATGCTGGCACACAAGGTGATGCTGGACTGGGCGGACGGGAGCTCTTCCAGGGCAGA | 648 |
| Db | 1314 | ----- | 1313 |
| Qy | 649 | TTTGATTATCTCTTACATCAACTGCAGGAGATGAACAGAGTGCCACCGAATGCAGC | 708 |
| Db | 1314 | -----AGGAGATGAACAGAGTGCCACCGAATGCAGC | 1346 |
| Qy | 709 | ATGCAAGACCTCATCTTCAGCTGTGGCTGAGCCCGAGCGGCTCTCCAGGAGCTCATC | 768 |
| Db | 1347 | ATGCAAGACCTCATCTTCAGCTGTGGCTGAGCCCGAGCGGCTCTCCAGGAGCTCATC | 1406 |
| Qy | 769 | CGAGTCCCGAGCGGCTCTTTTCATCATCGACGGCTTCGATAGCTCAAGCCTTCTTTC | 828 |
| Db | 1407 | CGAGTCCCGAGCGGCTCTTTTCATCATCGACGGCTTCGATAGCTCAAGCCTTCTTTC | 1466 |
| Qy | 829 | CACGATCTCAGGGACCTGGTGCTCTGTGGGAGGAGAAACGGCCCGAGCTGCTT | 888 |
| Db | 1467 | CACGATCTCAGGGACCTGGTGCTCTGTGGGAGGAGAAACGGCCCGAGCTGCTT | 1526 |
| Qy | 889 | CTTAACAGCTTAAATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGG | 948 |
| Db | 1527 | CTTAACAGCTTAAATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGG | 1586 |
| Qy | 949 | CCCAGGCTTTGGAGAGCTCCACGCTCTGTGGGACCGCCCGGCGATGTGGATCTTG | 1008 |
| Db | 1587 | CCCAGGCTTTGGAGAGCTCCACGCTCTGTGGGACCGCCCGGCGATGTGGATCTTG | 1646 |
| Qy | 1009 | GGCTTCTCTGAGGACGAAAGGAAGTACTTCTCAAGTATTTTCCCAATGCGAGAGCAG | 1068 |
| Db | 1647 | GGCTTCTCTGAGGACGAAAGGAAGTACTTCTCAAGTATTTTCCCAATGCGAGAGCAG | 1706 |
| Qy | 1069 | GCGGCGCAAGTCTTCAATTAAGTGGGACCAACGAGCTCTCTTCAACGATGCTTCGTC | 1128 |
| Db | 1707 | GCGGCGCAAGTCTTCAATTAAGTGGGACCAACGAGCTCTCTTCAACGATGCTTCGTC | 1766 |
| Qy | 1129 | CCCTGTGTGCTGGTGGTGTGTACTGCTCCAGCAGCAGCTGGAGGCTGGGGGCTG | 1188 |
| Db | 1767 | CCCTGTGTGCTGGTGGTGTGTACTGCTCCAGCAGCAGCTGGAGGCTGGGGGCTG | 1826 |

| | | | |
|----|------|--|------|
| Qy | 1189 | TTGAGACAGAGCTCCAGGACCAACCACTGCAGTGTACATGCTCTACCTGCTGAGTCTGATG | 1248 |
| Db | 1827 | TTGAGACAGAGCTCCAGGACCAACCACTGCAGTGTACATGCTCTACCTGCTGAGTCTGATG | 1886 |
| Qy | 1249 | CAACCCAAAGCGGGGGCCCCGCGCTCCAGCCCCCACCACCAACAGAGAGGGTGTGCTCC | 1308 |
| Db | 1887 | CAACCCAAAGCGGGGGCCCCGCGCTCCAGCCCCCACCACCAACAGAGAGGGTGTGCTCC | 1946 |
| Qy | 1309 | TTGCGGCGAGATGGGCTCTGGAATCAGAAATCCTATTGAGGAGCAGGACCTCCGGAAG | 1368 |
| Db | 1947 | TTGCGGCGAGATGGGCTCTGGAATCAGAAATCCTATTGAGGAGCAGGACCTCCGGAAG | 2006 |
| Qy | 1369 | CACGSCCTAGACGCGGGAGAGCGTCTCTGSCCTTCTCAACATGAACATCTTCCAGAAGGAC | 1428 |
| Db | 2007 | CACGSCCTAGACGCGGGAGAGCGTCTCTGSCCTTCTCAACATGAACATCTTCCAGAAGGAC | 2066 |
| Qy | 1429 | ATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGTTTCCAGGAATTTCTTTCGAGCT | 1488 |
| Db | 2067 | ATCAACTGTGAG----- | 2078 |
| Qy | 1489 | ATGTAATATATCTCTGGACGAGGGGGGCGGGCAGGCCCCAGACAGGACGTGACCAGG | 1548 |
| Db | 2079 | ----- | 2078 |
| Qy | 1549 | CTGTTGACCGAGTACGCGTTTCTGAAAGGAGCTTCTCTGGCACTCACCAGCGCTTCTG | 1608 |
| Db | 2079 | -----AGGAGCTTCTCTGGCACTCACCAGCGCTTCTG | 2111 |
| Qy | 1609 | TTTGGACTCTCTGAACGAGGAGACAGGAGCCACCTGGAGAAAGAGTCTCTGCTGGAAGGTC | 1668 |
| Db | 2112 | TTTGGACTCTCTGAACGAGGAGACAGGAGCCACCTGGAGAAAGAGTCTCTGCTGGAAGGTC | 2171 |
| Qy | 1669 | TCGCGCGACATCAAGATGGACCTGTTGCAAGTGGATCCAAAGCAAGCTCAGAGCGAGCGC | 1728 |
| Db | 2172 | TCGCGCGACATCAAGATGGACCTGTTGCAAGTGGATCCAAAGCAAGCTCAGAGCGAGCGC | 2231 |
| Qy | 1729 | TCCACCTCTCAGCAGGCGCTCTTGGAGTTCTTTCAGCTGCTTGTACGAGATCCAGGAGGAG | 1788 |
| Db | 2232 | TCCACCTCTCAGCAGGCGCTCTTGGAGTTCTTTCAGCTGCTTGTACGAGATCCAGGAGGAG | 2291 |
| Qy | 1789 | GAGTTTATCAGCAGGCGCTTACAGCCACTTTCAGGTGATCGTGTGTCAGCAACATTGCTCC | 1848 |
| Db | 2292 | GAGTTTATCAGCAGGCGCTTACAGCCACTTTCAGGTGATCGTGTGTCAGCAACATTGCTCC | 2351 |
| Qy | 1849 | AAGATGAGCAGATGCTCTCTGTTCTGTAAGCGCTGCGAGGCGCCAGCTGCTG | 1908 |
| Db | 2352 | AAGATGAGCAGATGCTCTCTGTTCTGTAAGCGCTGCGAGGCGCCAGCTGCTG | 2411 |
| Qy | 1909 | CACTTGTATGGCGCACCTTACAGCGCGGACGGGGAAGACCGCGAGGTGCTCCGCGAGGA | 1968 |
| Db | 2412 | CACTTGTATGGCGCACCTTACAGCGCGGACGGGGAAGACCGCGAGGTGCTCCGCGAGGA | 2471 |
| Qy | 1969 | GCGCACACGCTTTGTTGTCAGCTCAGACGAGAGGACCGTTCTGTGTCGAGCGCTACAGT | 2028 |
| Db | 2472 | GCGCACACGCTTTGTTGTCAGCTCAGACGAGAGGACCGTTCTGTGTCGAGCGCTACAGT | 2528 |
| Qy | 2029 | GAAATCTGCGAGCGGCGCTGTGCACCAATCCBAACCTGATAGAGTGTCTCTGTACCGGA | 2088 |
| Db | 2529 | GAAATCTGCGAGCGGCGCTGTGCACCAATCCBAACCTGATAGAGTGTCTCTGTACCGGA | 2588 |
| Qy | 2089 | AATGCCCTTGGGAGCGCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGC | 2148 |
| Db | 2589 | AATGCCCTTGGGAGCGCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGC | 2648 |
| Qy | 2149 | AACTTCAGAAACCTGA----- | 2164 |
| Db | 2649 | AACTTCAGAAACCTGAGGTAAAAATTTATCATATATATATATATATATATATATATATAT | 2708 |
| Qy | 2165 | ----- | 2164 |
| Db | 2709 | TTGGCCAGGATGATGAGTCAAGCTCAAGCTGTAATTCAGCACTTTTGGGAGGCGCCAGATGGGA | 2768 |

| | | | |
|----|------|--|------|
| QY | 2165 | ----- | 2164 |
| Db | 2769 | GGATCACTTGACCCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAACCCCATCTCTA | 2828 |
| QY | 2165 | ----- | 2164 |
| Db | 2829 | CTAAATAATCCAAATAGACCAGGCATGTGTGGCAACAGTCTGTAAAGCCAGCTACTCAGG | 2888 |
| QY | 2165 | -----GGCTGAAGAGGTGCC | 2179 |
| Db | 2889 | AGCCAAAGCAGAGGATTCCTCAACCCAGGAGGCAGAGGTTGTGGCTGAAGAGGTGCC | 2948 |
| QY | 2180 | GCATCTCAGCTCAGCCTGGAGGACCTCTCTGCAGCTCTCATGCCAATAGAAATTTGA | 2239 |
| Db | 2949 | GCATCTCCAGCTCAGCCTGCGAGGACCTCTCTGCAGCTCTCATAGCCAAATAAGAAATTTGA | 3008 |
| QY | 2240 | CAAGGATGGATCTCAGTGGCAACGGGTTGGATTCCAGGCATGATGCTGCTTTCGAGG | 2299 |
| Db | 3009 | CAAGGATGGATCTCAGTGGCAACGGGTTGGATTCCAGGCATGATGCTGCTTTCGAGG | 3068 |
| QY | 2300 | GCCTGGCGCATCCCCAGTGTGACGCTGCAGATGATTCAGTTGAGGAAGTGTACGCTGGAGT | 2359 |
| Db | 3069 | GCCTGGCGCATCCCAATGACGCTGCAGATGATTCAGTTGAGGAAGTGTACGCTGGAGT | 3128 |
| QY | 2360 | CCGGGGCTTCGACGAGATGGCTTCTGTGCTCGGCACCAACCCACATCTGTGTTGAGTTGG | 2419 |
| Db | 3129 | CCGGGGCTTCGACGAGATGGCTTCTGTGCTCGGCACCAACCCACATCTGTGTTGAGTTGG | 3188 |
| QY | 2420 | ACCTGACAGAAATGCACTGGAGATTTGGGCTGTAGGTTATATGCCAGGCACTGAGGC | 2479 |
| Db | 3189 | ACCTGACAGAAATGCACTGGAGATTTGGGCTGTAGGTTATATGCCAGGCACTGAGGC | 3248 |
| QY | 2480 | ACCCAGTCTGCAGACTACGACCTTTGTGGCTGAAGATCTGCCGCTCACTGTGCTGCCT | 2539 |
| Db | 3249 | ACCCAGTCTGCAGACTACGACCTTTGTGGCTGAAGATCTGCCGCTCACTGTGCTGCCT | 3308 |
| QY | 2540 | GTGACGAGCTGGCTCCAACTCTCAGTGTGAAACAGAGCCTGTAGAGAGCTGTGACCTGAGCC | 2599 |
| Db | 3309 | GTGACGAGCTGGCTCCAACTCTCAGTGTGAAACAGAGCCTGTAGAGAGCTGTGACCTGAGCC | 3368 |
| QY | 2600 | TGAATGAGCTGGGGACCTCGGGTGTGCTGTGCTGTGAGGGCCTCAGGCAATCCCACT | 2659 |
| Db | 3369 | TGAATGAGCTGGGGACCTCGGGTGTGCTGTGCTGTGAGGGCCTCAGGCAATCCCACT | 3428 |
| QY | 2660 | GCAAGCTCCAGACCTCGGGTTGGC | 2685 |
| Db | 3429 | GCAAGCTCCAGACCTCGGGTGATC | 3454 |

```

RESULT 12
US-10-183-770-1
; Sequence 1, Application US/10183770
; Publication No. US20030180812A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; FILE REFERENCE: D0067A CIP
; CURRENT APPLICATION NUMBER: US/10/183,770
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 10/028,374
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4931
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1320)..(2666)

```

| | | | | | | | | | |
|---|------|--|------|--|--|--|--|--|--|
| ; OTHER INFORMATION: | | | | | | | | | |
| US-10-183-770-1 | | | | | | | | | |
| Query Match 52.0%; Score 1616; DB 16; Length 4931; | | | | | | | | | |
| Best Local Similarity 81.0%; Pred. No. 0; | | | | | | | | | |
| Matches 2160; Conservative 0; Mismatches 5; Indels 501; Gaps 4; | | | | | | | | | |
| Qy | 289 | GATCCCAGGAACTACAGGACATGTGCGCAGGAAATTCGGCTCATGGAAGACCGC | 348 | | | | | | |
| Db | 1021 | GATCCCAGGAACTACAGGACATGTGCCAGGAAATTCGGCTCATGGAAGACCGC | 1080 | | | | | | |
| Qy | 349 | AATCGCGCTAGGGGAATGTGTCAACTCAGCCACCGGTACACCCGGCTCTGCTGTTG | 408 | | | | | | |
| Db | 1081 | AATCGCGCTAGGGGAATGTGTCAACTCAGCCACCGGTACACCCGGCTCTGCTGTTG | 1140 | | | | | | |
| Qy | 409 | AAGGAGACTCAAAACCCCATGCAAGTCCAGCAGCAGCTTCTGGACACAGGCCGGGACAC | 468 | | | | | | |
| Db | 1141 | AAGGAGACTCAAAACCCCATGCAAGTCCAGCAGCAGCTTCTGGACACAGGCCGGGACAC | 1200 | | | | | | |
| Qy | 469 | GGGAGGACCGTGGGACACAGAGCTAGCCCATCAAGATAGAGACCCCTCTTTGAGCAGAC | 528 | | | | | | |
| Db | 1201 | GGGAGGACCGTGGGACACAGAGCTAGCCCATCAAGATAGAGACCCCTCTTTGAGCAGAC | 1260 | | | | | | |
| Qy | 529 | GAGGAGCGCCCGAGGCAACGCGCACCGTGTCTATGCAAGCGCGCAGGATAGGCAAG | 588 | | | | | | |
| Db | 1261 | GAGGAGCGCCCGAGGCAACGCGCACCGTGTGTATGCAAGCGCGCAGGAT----- | 1313 | | | | | | |
| Qy | 589 | TCCATGCTGGCACACAAGGTGATGTGGACTGGGCGGACGGAAAGCTTTTCCAGGCAGA | 648 | | | | | | |
| Db | 1314 | ----- | 1313 | | | | | | |
| Qy | 649 | TTTGATTTATCTCTTTACATCAACTGGAGGAGATGAACAGAGTGCACGGNATGCCAGC | 708 | | | | | | |
| Db | 1314 | -----AGGGAGATGAACAGAGTGCACGGNATGCCAGC | 1346 | | | | | | |
| Qy | 709 | ATGCAAGACCTCATCTTCAGCTGTCGCCCTGAGCCACGCGCGCTCTCCAGGAGCTCATC | 768 | | | | | | |
| Db | 1347 | ATGCAAGACCTCATCTTCAGCTGTCGCCCTGAGCCACGCGCGCTCTCCAGGAGCTCATC | 1406 | | | | | | |
| Qy | 769 | CGAGTTCGCGAGCGCTCTCTTTTCATCATCGACGGCTTGATGAGCTCAAGCTTCCTTC | 828 | | | | | | |
| Db | 1407 | CGAGTTCGCGAGCGCTCTCTTTTCATCATCGACGGCTTGATGAGCTCAAGCTTCCTTC | 1466 | | | | | | |
| Qy | 829 | CACGATCTCAGGACCCCTGGTGCTCTGTGGGAGGAGAAACGGCCACGGAGCTGCTT | 888 | | | | | | |
| Db | 1467 | CACGATCTCAGGAGCCCTGGTGCTCTGTGGGAGGAGAAACGGCCACGGAGCTGCTT | 1526 | | | | | | |
| Qy | 889 | CTTAAACAGCTTAAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACCACGG | 948 | | | | | | |
| Db | 1527 | CTTAAACAGCTTAAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACCACGG | 1586 | | | | | | |
| Qy | 949 | CCACGGCTTTGGAGAGCTCCACCGTCTCTGGAGCACCCACGGCATGTGGAGATCTCTG | 1008 | | | | | | |
| Db | 1587 | CCACGGCTTTGGAGAGCTCCACCGTCTCTGGAGCACCCACGGCATGTGGAGATCTCTG | 1646 | | | | | | |
| Qy | 1009 | GGCTTCTCTGAGGACAGAAAGGAGAAATATCTTCTACAAGTATTTCCAATGAGAGCAG | 1068 | | | | | | |
| Db | 1647 | GGCTTCTCTGAGGACAGAAAGGAGAAATATCTTCTACAAGTATTTCCAATGAGAGCAG | 1706 | | | | | | |
| Qy | 1069 | GGGGGCCAAGTCTCAATTAGCTGAGGGAACAACGAGCTCTCTTCAACCATGTGCTTCGTC | 1128 | | | | | | |
| Db | 1707 | GGGGGCCAAGTCTCAATTAGCTGAGGGAACAACGAGCTCTCTTCAACCATGTGCTTCGTC | 1766 | | | | | | |
| Qy | 1129 | CCCTGTGTGTGGTGGTGTGTACTGCTCCAGCAGCAGCTGGAGGGTGGGGGCTG | 1188 | | | | | | |
| Db | 1767 | CCCTGTGTGTGGTGGTGTGTGTACTGCTCCAGCAGCAGCTGGAGGGTGGGGGCTG | 1826 | | | | | | |
| Qy | 1189 | TTGAGACAGAGTCCAGGACCAACATGTCAGTGTACATGCTCTACTGCTGAGTCTCATG | 1248 | | | | | | |
| Db | 1827 | TTGAGACAGAGTCCAGGACCAACATGTCAGTGTACATGCTCTACTGCTGAGTCTCATG | 1886 | | | | | | |
| Qy | 1249 | CAACCCAGCGGGGGCCCCCGCGCTCCAGCCCCCACCAACACAGAGAGGGTCTGTCTCC | 1308 | | | | | | |

Db 1887 CAACCAAGCCGGGGCCCCGGCCTCCAGCCCCCAACCAACAGAGAGGGTTGTGCTCC 1946
QY 1309 TTGGCGGAGATGGGCTCTGGAATCAGAAATCTATTATTGAGGAGCAGGACCTCCGGAAG 1368
Db 1947 TTGGCGGAGATGGGCTCTGGAATCAGAAATCTATTATTGAGGAGCAGGACCTCCGGAAG 2006
QY 1369 CACGGCCTAGACGGGGAGAGCGTCTCTGCTCTCTCAACATGAACATCTTCCAGAGGAC 1428
Db 2007 CACGGCCTAGACGGGGAGAGCGTCTCTGCTCTCTCAACATGAACATCTTCCAGAGGAC 2066
QY 1429 ATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAAGTTCCAGGAATCTTTTGACGCT 1488
Db 2057 ATCAACTGTGAG----- 2078
QY 1489 ATGTAATAATCTCTGGAACGAGGGGAGCGGGCAGGCCAGACAGGACGTGACCAAG 1548
Db 2079 ----- 2078
QY 1549 CTGTTGACCGAGTAGCGCTTTTCTGAAGGAGCTTCTTGGCACTCACCAGCGGCTTCTG 1608
Db 2079 -----AGGAGCTTCTTGGCACTCACCAGCGGCTTCTG 2111
QY 1609 TTTGGACTCTGAACGAGGACAGGAGCCACCTGGAGAAGAGTCTCTGCTGGAAGTC 1668
Db 2112 TTTGGACTCTGAACGAGGACAGGAGCCACCTGGAGAAGAGTCTCTGCTGGAAGTC 2171
QY 1669 TCGCGGCACATCAAGATGGAACCTGTTGAGTGGATCCAAAGCAAGCTCAGAGGACGGC 1728
Db 2172 TCGCGGCACATCAAGATGGAACCTGTTGAGTGGATCCAAAGCAAGCTCAGAGGACGGC 2231
QY 1729 TCCACCTTGACGAGGCTCTTGGAGTCTTACGATCTGTTGACGAGATCCAGAGGAG 1788
Db 2232 TCCACCTTGACGAGGCTCTTGGAGTCTTACGATCTGTTGACGAGATCCAGAGGAG 2291
QY 1789 GAGTTTATCCAGCAGGCGCTCAGCCACTTCCAGGTGATCGTGTGAGCAAAATGTCCTCC 1848
Db 2292 GAGTTTATCCAGCAGGCGCTCAGCCACTTCCAGGTGATCGTGTGAGCAAAATGTCCTCC 2351
QY 1849 AAGATGAGACATGGTCTCTCGTTCTGTCTGAGCGCTCAGAGGCGCCAGGTGCTG 1908
Db 2352 AAGATGAGACATGGTCTCTCGTTCTGTCTGAGCGCTCAGAGGCGCCAGGTGCTG 2411
QY 1909 CACTTGTATGGCGCACCTCAGCGCGGACGGGGAAGACCGGCGGAGGTGCTCCGCGGA 1968
Db 2412 CACTTGTATGGCGCACCTCAGCGCGGACGGGGAAGACCGGCGGAGGTGCTCCGCGGA 2471
QY 1969 GCGCACGCTGTTGTGCTCAGTCCAGCAGAGGACCGTCTGCTGAGCGCTACAGT 2028
Db 2472 GCGCACGCTGTTGTGCTCAGT---ACCAGAGGACCGTCTGCTGAGCGCTACAGT 2528
QY 2029 GAACATCTGGCAGCGGCGCTGTGCAACCAATCCAACTGTATAGAGCTGCTCTGTACCGA 2088
Db 2529 GAACATCTGGCAGCGGCGCTGTGCAACCAATCCAACTGTATAGAGCTGCTCTGTACCGA 2588
QY 2089 AATGCCCTGGGAGCGCGGGGGTGAAGTGTCTGTGAAGACTCAGACACCCCACTGTC 2148
Db 2589 AATGCCCTGGGAGCGCGGGGGTGAAGTGTCTGTGAAGACTCAGACACCCCACTGTC 2648
QY 2149 AAACCTCAGAACCTGA----- 2164
Db 2649 AAACCTCAGAACCTGAGGTAATAATATATATATATATATATATATATATATATATATAT 2708
QY 2165 ----- 2164
Db 2709 TTGGCCAGGTATGATGGCTCAGCGCTGTAATTCAGCACTTTGGGAGGCCAGATGGGA 2768
QY 2165 ----- 2164
Db 2769 GGATCACTTGACCCAGGAGTTCAAGACCAAGCGCTGGCCAAACATGGTGAACCCCATCTCTA 2828
QY 2165 ----- 2164
Db 2829 CTAATAATAACAAATAGCAGGATGGTGGCACAGCTGTGTAAAGCCCACTACTCAGG 2888

QY 2165 -----GGCTGAAGAGGTGCC 2179
Db 2889 AGGCAAGGACGAGGAGTTGCTTCAACCCAGAGGACAGAGGTGTGTGGCTGAAGAGGTGCC 2948
QY 2180 GCATCTCCAGCTCAGCCTGGGAGACCTCTCTGAGCTCTCATAGCCCAATGAATTTGA 2239
Db 2949 GCATCTCCAGCTCAGCCTGGGAGACCTCTCTGAGCTCTCATAGCCCAATGAATTTGA 3008
QY 2240 CAAGGATGGATCTCAGTGGCAACGGCGTTGGATTCCCAGGACATGATGCTGCTTTGCCGAGG 2299
Db 3009 CAAGGATGGATCTCAGTGGCAACGGCGTTGGATTCCCAGGACATGATGCTGCTTTGCCGAGG 3068
QY 2300 GCCTGCGGCATCCCAGCTGCAAGGCTGCAGATGATTCAATTGAGGAAGTGTGAGCTGGAGT 2359
Db 3069 GCCTGCGGCATCCCAGCTGCAAGGCTGCAGATGATTCAATTGAGGAAGTGTGAGCTGGAGT 3128
QY 2360 CCGGGGCTTGTCCAGGATGGCTTCTGTCTGCGCACCAACCCACATCTGTTGAGTTGG 2419
Db 3129 CCGGGGCTTGTCCAGGATGGCTTCTGTCTGCGCACCAACCCACATCTGTTGAGTTGG 3188
QY 2420 ACCTGACAGGAATGCACCTGAGGATTTGGGCTGAGGTTACTATGCCAGGACTGAGGC 2479
Db 3189 ACCTGACAGGAATGCACCTGAGGATTTGGGCTGAGGTTACTATGCCAGGACTGAGGC 3248
QY 2480 ACCAGTCTGCAGACTACGGAATTGTTGGCTGGAAGATCTGCGGCTCTCACTGCTGCTGCT 2539
Db 3249 ACCAGTCTGCAGACTACGGAATTGTTGGCTGGAAGATCTGCGGCTCTCACTGCTGCTGCT 3308
QY 2540 GTGACGAGCTGGGCTCAACTCTCAGTGTGAACAGAGCTGAGAGCTGAGACCTGAGGC 2599
Db 3309 GTGACGAGCTGGGCTCAACTCTCAGTGTGAACAGAGCTGAGAGCTGAGACCTGAGGC 3368
QY 2600 TGAATGAGCTGGGGGACCTCGGGGTGCTGCTGCTGCTGAGGGGCTCAGGACATCCACGT 2659
Db 3369 TGAATGAGCTGGGGGACCTCGGGGTGCTGCTGCTGCTGAGGGGCTCAGGACATCCACGT 3428
QY 2660 GCAAGCTCCAGACCTCGGTTGGGC 2685
Db 3429 GCAAGCTCCAGACCTCGGTTGAGTC 3454

RESULT 13

US-10-162-335-23

; Sequence 23, Application US/10162335

; Publication No. US20040009480A1

; GENERAL INFORMATION:

; APPLICANT: Anderson, David W.

; APPLICANT: Baumgartner, Jason C.

; APPLICANT: Boldog, Ferenc I.

; APPLICANT: Casman, Stacie J.

; APPLICANT: Edinger, Shlomit R.

; APPLICANT: Gangolli, Esha A.

; APPLICANT: Gerlach, Valerie

; APPLICANT: Gorman, Linda

; APPLICANT: Guo, Xiaojia (Sasha)

; APPLICANT: Hjalt, Tord

; APPLICANT: Kekuda, Rameeh

; APPLICANT: Li, Li

; APPLICANT: MacDougall, John R.

; APPLICANT: Malyankar, Uriel M.

; APPLICANT: Millet, Isabelle

; APPLICANT: Padigar, Muralidhara

; APPLICANT: Patturajan, Meera

; APPLICANT: Pena, Carol E. A.

; APPLICANT: Rastelli, Luca

; APPLICANT: Shimkets, Richard A.

; APPLICANT: Spytek, Kimberly A.

; APPLICANT: Vernet, Corine A. M.

; APPLICANT: Voos, Edward Z.

; APPLICANT: Zerhusen, Bryan D.

; TITLE OF INVENTION: Therapeutic Polypeptides, Nucleic Acids Encoding Same, and Metho

[illegible]

RESULT 14

RESOL 14
US-09-799-983-1

00 03 725 203-1
: Sequence 1, Application US/09799983

; Sequence 1, Application US/
; Patent No. US20010029033A1

FACEPIC NO. US2001002
: GENERAL INFORMATION:

APPLICANT: Shami. Paul

APPLICANT: SHAMI, PAUL

APPLICANT: PARKER, CHARLES
TITLE OF INVENTION: NOVEL GENE PRO LIPID

| | | |
|---|---------------------|-----------|
| ; | TITLE OF INVENTION: | NOVEL |
| : | FILE REFERENCE: | 1337 2 51 |

```

; FILE REFERENCE: 1321.2.51
; CURRENT APPLICATION NUMBER: 108/00/700 000

```

;; CURRENT APPLICATION NUMBER: US/09/799,983
: CURRENT FILING DATE: 2001 03 05

; CURRENT FILING DATE: 2001-03-06
 ; PRIOR APPLICATION NUMBER: US 2001/000 000

US 60/186,971

; PRIOR FILING DATE: 2000-03-06

; NUMBER OF SEQ ID NOS: 24

```

; SOFTWARE: PatentIn version 3.0

```

; SEQ ID NO 1

; LENGTH: 2682

; TYPE: DNA

; ORGANISM: Homo sapiens

Query Match 30.5%: Score 947.6: DB 9: Length 2682:

quality match
 Best Local Similarity
 99.6%: Pred. No. 1.3e-262:
 30.3%, score 347.0, DB 3;
 Tennyson 2002;

```

Best Local Similarity 95.6%; Freq. NO. 1.3E-262;
Matches 950: Conservative 0: Mismatches 4: Indels 0: Gaps 0:

```

| | | | |
|------|----|--|------|
| 2155 | Qy | CAGAACTGTAGGCTGAAGAGGTGCCGCAATCTTCAGCTCAGCCTGCGAGGACCTCTCTGCA | 2214 |
| 1407 | Db | CAGAGGTGTGGCTGAAGAGGTGCGCATCTCAGCTCAGCCTGCGAGGACCTCTCTGCA | 1466 |
| 2215 | Qy | GCTCTCATAGCCCAATAGAATTTGACAAGATGGATCTCAGTGGCAACGCCGTTGGATTTC | 2274 |
| 1467 | Db | GCTCTCATAGCCCAATAGAATTTGACAAGATGGAATCTCAGTGGCAACGCCGTTGGATTTC | 1526 |
| 2275 | Qy | CCAGGCATGATGCTGCTTTGCGAGGGCTCGCGCATCCCAAGTCAGGCTGCAGATGCAATT | 2334 |
| 1527 | Db | CCAGGCATGATGCTGCTTTGCGAGGGCTCGCGCATCCCAAGTCAGGCTGCAGATGCAATT | 1586 |
| 2335 | Qy | CAGTTGAGGAAGTGTCAAGCTCGGAGTCGGGGCTTTGTCAAGAGATGGCTTCTGTGCTCGGC | 2394 |
| 1587 | Db | CAGTTGAGGAAGTGTCAAGCTCGGAGTCGGGGCTTTGTCAAGAGATGGCTTCTGTGCTCGGC | 1646 |
| 2395 | Qy | ACCAACCCACATCTGGTTGATTTGACCTGACAGGAATGCACTGGAGGATTTGGGCGCTG | 2454 |
| 1647 | Db | ACCAACCCACATCTGGTTGATTTGACCTGACAGGAATGCACTGGAGGATTTGGGCGCTG | 1706 |
| 2455 | Qy | AGGTTACTATGCCAGGACTGAGGCCACCAAGTCTGCAGACTACGGACTTTGTGGCTGAAG | 2514 |
| 1707 | Db | AGGTTACTATGCCAGGACTGAGGCCACCAAGTCTGCAGACTACGGACTTTGTGGCTGAAG | 1766 |
| 2515 | Qy | ATCTGCCGCTCACTGCTGCTGTGTGACGAGCTGGCCCTCAACTCTCAGTGTGAACCCAG | 2574 |
| 1767 | Db | ATCTGCCGCTCACTGCTGCTGTGTGACGAGCTGGCCCTCAACTCTCAGTGTGAACCCAG | 1826 |
| 2575 | Qy | AGCCTGAGAGAGCTGGAACCTGAGCCTGAATGAGCTGGGGACCTCGGGGTGCTGCTGCTG | 2634 |
| 1827 | Db | AGCCTGAGAGAGCTGGAACCTGAGCCTGAATGAGCTGGGGACCTCGGGGTGCTGCTGCTG | 1886 |
| 2635 | Qy | TGTGAGGGGCTCAGGCATCCACGTGCAAGCTCAAGACCTCGGGTTGGGCACTGTCGGG | 2694 |
| 1887 | Db | TGTGAGGGGCTCAGGCATCCACGTGCAAGCTCAAGACCTCGGGTTGGGCACTGTCGGG | 1946 |

| | | | |
|----|------|--|------|
| Qy | 461 | GGGACACGCGAGGACCGTGGGACACGAGCTAGCCCCATCAAGATPAGAGACCCCTTTTG | 520 |
| Db | 568 | -----AAGACCAAGAGCTGTGAGAGCCCGTGAGTCCCATTAAGATGAGATTGCTGTTTG | 622 |
| Qy | 521 | AGCCAGACGAGGAGCGCCCGAGACCGCGCACCGTGTGTCATGACGAGCGCGGAGGGA | 580 |
| Db | 623 | ACCCGATGATGAGCATTTGAGCCCTGTGCACACCGTGTGTTCAGGGGGGGCGGAGGA | 682 |
| Qy | 581 | TAGCAAGTCCATGCTGGGCACACAGGTGATGCTGACTGGGCGGACGGGAAGCTCTTCC | 640 |
| Db | 683 | TTGGGAANAACATCTGCGCCAGAGAGATGATGTGACTGGGATCGGGGACACTCTACC | 742 |
| Qy | 641 | AAGCAGATTGATTATCTCTTCAATCAATGCAAGGAGAGATGAACGAGATGCCACGG | 700 |
| Db | 743 | AAGACAGTTTGACTATCTGTCTATATCACTGTGCGGAGGTGAGCC---TTGTGACAC | 799 |
| Qy | 701 | AATCAGCATGCAAGACCTCATCTTCAGCTGCTGGCTGAGCCGAGCGGCTCTCCAGG | 760 |
| Db | 800 | AGAGAGCCTGGGGACCTGATCATGAGCTGCTGCCCGACCCAAACCCACCATCCACA | 859 |
| Qy | 761 | AGCTCATCCGAGTTCCGAGCGCTCCTTTTTCATCATGACGGCTTCGATGAGCTCAAGC | 820 |
| Db | 860 | AGATCGTGAGAAACCTCCAGATCTCTTCTCATGAGCGCTTCGATGAGCTGCAAG | 919 |
| Qy | 821 | CTTCTTTCCAGTCTTOAGGACCCCTGGTGCCTCTGCTGGGAGGAGAAACGCCACCG | 880 |
| Db | 920 | GTGCTTTGACGAGCACATAGAACCCGCTCTGCACTGACTGGCAGAAAGCGCGGGAG | 979 |
| Qy | 881 | AGCTGCTTCTTAAACAGCTTAATTCGAGAGAGCTCTCCCTGAGCTATCTTTGCTCATCA | 940 |
| Db | 980 | ACATTTCTCCTGAGCAGCCTCATAGAAAGAGCTCTTCCGAGCCCTCTGCTCATCA | 1039 |
| Qy | 941 | CCACAGCGCCACCGCTTTGAGAGCTTCAACCTGCTGGAGCAACCCAGGAGATGG | 1000 |
| Db | 1040 | CCAGAGACCTGTGSCCTGGAGAACTGACGACCTTGTCTGGACCATCTCGGCATGG | 1099 |
| Qy | 1001 | AGATCTGGGCTTCTGAGGAGAGAGAGAGTACTTCTACAAGTATTTCCACAAATG | 1060 |
| Db | 1100 | AGATCTGGGTTTCTCGAGGCCAAAGGAAAGAGTACTTCTCAAGTATTTCTCTGATG | 1159 |
| Qy | 1061 | CAGACAGCGGGCCAAAGTCTTCAATTAAGTGGGAGCAACGAGCCTCTCTTCAACATGT | 1120 |
| Db | 1160 | AGGCCAGACGAGGACCTTCAGTCTGATTCAGGAGAACGAGTCTCTTCAACATGT | 1219 |
| Qy | 1121 | GCTTCGTCCTGCTGTGCTGGTGTGTGTAACCTGCTCCAGCAGCAGCTGAGGGTG | 1180 |
| Db | 1220 | GCTTCATCCCCCTGCTGCTGATCGTGTGCACTGGACTGAAACAGCAGATGAGAGTG | 1279 |
| Qy | 1181 | GGGGCTGTTGAGACAGAGCTCCAGGACCAACCTGCACTGATGTACATGCTCTACCTGTGA | 1240 |
| Db | 1280 | GCAAGAGCCTTGCCAGACATTAAGACAACAACGCGGTGTACGTCTTCTTCTTCCCA | 1339 |
| Qy | 1241 | GTCTGATGCAACCCAGCGGGGGCCCCGCGCTCCAGCCCCCAACCCAGCAGAGAGGT | 1300 |
| Db | 1340 | GTTTGTGTCAGCCCCGGGAGGAGCCAGGAGCAGGCTCTGCGCCCCACCTCTGGGGC | 1399 |
| Qy | 1301 | TGTGCTCTTGGGGCAGATGGGCTCTGGAATCAGAAATCCTATTTGAGGAGCAGGACC | 1360 |
| Db | 1400 | TCTGCTCTTTGGCTGAGATGAATCTGGAACCAAGAAATCTGTTGAGGAGTCCGACC | 1459 |
| Qy | 1361 | TCCGAAGCAGCGCTTAGACGGGGAAGAGCTCTGCTGCTTCTTCAACATGAACATCTTCC | 1420 |
| Db | 1460 | TCAGGAATCATGGACTGAGAGGGCGGATGTGTCTGCTTCTCTAGGATGAACCTGTTC | 1519 |
| Qy | 1421 | AGAAGACATCAACTGTGAGAGGTACTACAGCTTCACTGATGATTTCCAGGATTTCT | 1480 |
| Db | 1520 | AAAAGGAAATGAGCTGCGGAGATTTCTACAGCTTCACTCCACATGCTTTCCAGGAGTCT | 1579 |
| Qy | 1481 | TTGCACTGATGACTATATCTTGAGAGGGGGAGGGCGG----- | 1520 |
| Db | 1580 | TTCCGCCCATGTACTACCTGCTGGAGAGGAAAGGAGGAGGACGTTCCAGGGA | 1639 |
| Qy | 1521 | -----GGCAGGCCAGACACGAGCCTGACCCAGCTGTTGACCGAGTACCGCTTTCTG | 1573 |
| Db | 1640 | GTCCGTTTGAAGCTTCCAGCCGAGACCTGACAGTCTTCTGGAATACTATGGAATACTG | 1699 |
| Qy | 1574 | AAAGGAGCTTCTCGCACTCACACGCGCTTCTGTTTGGACTCCTGAACGAGGAGACCA | 1633 |
| Db | 1700 | AAAAGGGTATTTGATTTTGTGTGCTGTTTCTCTTGGCTGGTAAACGAGGAGGA | 1759 |
| Qy | 1634 | GGAGCCACTGGAGAGAGTCTCTGCTGGAAGTCTCGCCGACATCAAGATGACACCTGT | 1693 |
| Db | 1760 | CCTCCTACTTGGAGAGAAATTAAGTTGCAAGATCTCTCAGCAATCAGGCTGGAGTGC | 1819 |
| Qy | 1694 | TGCACTGATCCAAAGCAAAGCTCAGAGCGAGCGCTCCACCTGCGAGGAGCTCTTGG | 1753 |
| Db | 1820 | TGAATGATTAAGTGAAGCCAAAGCTAAAGCTGCAGATCCAGCCAGCAGCTGG | 1879 |
| Qy | 1754 | AGTTCTCAGCTGCTGTGACAGATCCAGGAGGAGGATTTATCCAGCAGGCCCTGAGCC | 1813 |
| Db | 1880 | AATGTTCTACTGTTGTACGAGATGCAGGAGGAGGACTTCGTGCAAGGGCCATGGA | 1939 |
| Qy | 1814 | ACTTCCAGGTGATCGTGTGACCAACATTCCTCCAAAGATGGAGCAATGCTCTCTCGT | 1873 |
| Db | 1940 | ATTTCCCCCAAGATTGAG---ATCAATCTCTCCACAGAAATGACCAATGTTCTTCT | 1996 |
| Qy | 1874 | TCTGTCTGAAGCGCTGCAGGAGCGCCAGGTGCTGCACTTGTATGGGCCACCTACAGCG | 1933 |
| Db | 1997 | TTTGCAATTGAGAACTGTCTATCGGGTGGAGTCACTGTCTCCCTGGGGTTTCTCATAAATGC | 2056 |
| Qy | 1934 | CSGACGGGGAGA-----CCGCGCGAGGTGC | 1959 |
| Db | 2057 | CCAAGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG | 2116 |
| Qy | 1960 | TCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG | 2019 |
| Db | 2117 | TCCAGGCTCTCTCATGCTGCTGCTTC-TCATGATTGCTGAACAGGAGGAGGAGGAGG | 2175 |
| Qy | 2020 | GCTCAGTGAACATCTGGCAGCGCCCTGTGACCAATCCAAACCTCATAGAGCTGTCT | 2079 |
| Db | 2176 | AGTTTTTGGCGGGCTCTTTTTCAGTTCTGAGCAGCAGGAGGAGGAGGAGGAGGAGG | 2235 |
| Qy | 2080 | CTGTACCGAAATGCTGCGGAGCGGGGGTGAAGCTGCTCTGTCAGGAGCTCAGACAC | 2139 |
| Db | 2236 | CTCAGTGACATTTCTTGGGGGACCCAGGGATGAGAGTGTGTGTGAACCGTCCAGCAT | 2295 |
| Qy | 2140 | CCCACTGCAAACTTCAGAACTGAGGCTGAAGAGTGGCGCATCTTCAGCTCAGCCTGC | 2199 |
| Db | 2296 | CCTGGCTGAACATTCGAGATTGTGTTGGGGCGCTGTGGCTCTCGCATGAGTGTCTGC | 2355 |
| Qy | 2200 | GAGGACCTCTCTGACCTCTCATAGCCAATAAGAAATTTGACAAGGATGATCTCAGTGGC | 2259 |
| Db | 2356 | TTCCGACATCTCTTGGTCTCTCAGCAGCAACCAAGAGCTGGTGGAGCTGGACCTGAGTGAC | 2415 |
| Qy | 2260 | AAGCGGCTTGGATTCACAGGATGATGCTGTTTGGAGGCGCTGGGGATCCCAAGTGC | 2319 |
| Db | 2416 | AAGCGCTCGGTGACTTCGGAATCAGACTTCTGTGTGGGACTGAAGCAGCTGTTGTGC | 2475 |
| Qy | 2320 | AGGCTGAGATGATTCAGTTGAGGAGTGTACAGTGGAGTCCGGGCTTGTGAGGAGATG | 2379 |
| Db | 2476 | AATCTGAAGAAGCTCTGGTTGGTCTGCTGCTCCTCATCAGCATGTTGTGAGGATCTT | 2535 |
| Qy | 2380 | GCTTCTGTGCTCGGACCAACCCACATCTGTTTGAAGTGGAGCTCTGACAGGAAATGAC | 2439 |
| Db | 2536 | GCATCAGTATTGAGCAGCAGCCATTCCTTGAACAGACTCTATGTGGGGAGGATGCTTGC | 2595 |
| Qy | 2440 | GAGGATTTGGGCTGAGGTTTATATGCAAGGAGCTGAGGACCCAGTCTGCACTACCG | 2499 |
| Db | 2596 | GGAGACTCAGGAGTGCAGATTTTATGTGAAAAAGCAAGAAATCCACAGTGAACCTGCAG | 2655 |
| Qy | 2500 | ACTTTGGCTGAGATCTCGCCCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 2559 |
| Db | 2656 | AAACTGGGGTGGTGAATTCGTCCTTACGTGCTGTGTTGTTGTTGTTGTTGTTGTTGTTG | 2715 |
| Qy | 2560 | CTCAGTGTGAACAGAGCCTGAGAGAGCTGGAACCTGAGCCTGATGAGCTGGGGGACCTC | 2619 |

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 8, 2005, 15:58:31 ; Search time 8928 Seconds
(without alignments)
13250.868 Million cell updates/sec

Title: US-10-781-294-23

Perfect score: 3108

Sequence: 1 atgtctgaacccgagcag.....atttgacattgctgtcta 3108

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hc:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_ges1:*
- 9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 762.4 | 24.5 | 942 | 5 | BQ957631 |
| 2 | 698 | 22.5 | 917 | 5 | BQ920369 |
| 3 | 660.2 | 21.2 | 937 | 5 | BQ900330 |
| 4 | 539.8 | 17.4 | 571 | 2 | AM673661 |
| 5 | 518.8 | 16.7 | 664 | 9 | AG067278 |
| 6 | 494.2 | 15.9 | 568 | 7 | AW956628 |
| 7 | 493.8 | 15.9 | 669 | 7 | CV030364 |
| 8 | 465.2 | 15.0 | 496 | 8 | AQ726243 |
| 9 | 463.8 | 14.9 | 499 | 8 | AQ814983 |
| 10 | 426 | 13.7 | 764 | 4 | B1911853 |
| 11 | 424.4 | 13.7 | 449 | 8 | AQ443134 |
| 12 | 398 | 12.8 | 4123 | 3 | HSM000983 |
| 13 | 395 | 12.7 | 396 | 8 | AQ094928 |
| 14 | 377 | 12.1 | 3359 | 3 | BC021272 |
| 15 | 377 | 12.1 | 3360 | 3 | BC012789 |
| 16 | 366.4 | 11.8 | 788 | 4 | EM043219 |
| 17 | 354.6 | 11.4 | 983 | 5 | BQ710940 |
| 18 | 349.8 | 11.3 | 581 | 9 | AG161279 |
| 19 | 343.2 | 11.0 | 3475 | 3 | AK087774 |
| 20 | 338.2 | 10.9 | 2218 | 3 | AF054176 |
| 21 | 330.8 | 10.6 | 643 | 4 | B1519910 |
| 22 | 287.6 | 9.3 | 490 | 9 | C6648167 |
| 23 | 282 | 9.1 | 496 | 9 | C6663401 |
| 24 | 281 | 9.0 | 772 | 6 | CB947742 |

| | | | | | | |
|----|-------|-----|------|---|----------|-----------|
| 25 | 280 | 9.0 | 2350 | 3 | CR622836 | full-leng |
| 26 | 276.6 | 8.9 | 459 | 8 | AQ716650 | HS_5447_B |
| 27 | 272.6 | 8.8 | 1371 | 9 | AY407367 | Homo sapi |
| 28 | 272.6 | 8.8 | 1507 | 3 | CR613396 | full-leng |
| 29 | 272.6 | 8.8 | 1548 | 3 | CR604056 | full-leng |
| 30 | 272.6 | 8.8 | 1612 | 3 | CR592774 | full-leng |
| 31 | 272.6 | 8.8 | 1623 | 3 | CR594328 | full-leng |
| 32 | 272.6 | 8.8 | 1630 | 3 | CR604995 | full-leng |
| 33 | 272.6 | 8.8 | 1647 | 3 | CR606240 | full-leng |
| 34 | 272.6 | 8.8 | 1650 | 3 | CR597292 | full-leng |
| 35 | 272.6 | 8.8 | 1653 | 3 | CR625171 | full-leng |
| 36 | 272.6 | 8.8 | 1657 | 3 | CR590338 | full-leng |
| 37 | 272.6 | 8.8 | 1661 | 3 | CR614192 | full-leng |
| 38 | 272.6 | 8.8 | 1662 | 3 | CR620826 | full-leng |
| 39 | 272.6 | 8.8 | 1664 | 3 | CR600355 | full-leng |
| 40 | 272.6 | 8.8 | 1678 | 3 | CR600411 | full-leng |
| 41 | 272.6 | 8.8 | 1679 | 3 | CR613915 | full-leng |
| 42 | 272.6 | 8.8 | 1680 | 3 | CR624221 | full-leng |
| 43 | 272.6 | 8.8 | 1684 | 3 | CR602999 | full-leng |
| 44 | 272.6 | 8.8 | 1685 | 3 | CR597816 | full-leng |
| 45 | 272.6 | 8.8 | 1687 | 3 | CR608218 | full-leng |

ALIGNMENTS

RESULT 1
BQ957631
LOCUS
DEFINITION AGENCOURT_8775651 NIH_MGC_40 Homo sapiens cdna clone IMAGE:6383428
5', mRNA sequence.
ACCESSION BQ957631
VERSION BQ957631.1 GI:22373109
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 942)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2576 row: n column: 05
High quality sequence stop: 453.
Location/Qualifiers
1. .942
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6383428"
/issue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI, cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match

24.5%; Score 762.4; DB 5; Length 942;


```
QY 1924 ACCTACAGCGGAGCGGGAAGACCGCGGAGGTGCTCCGAGAGGCGCACACGCTGTG 1983
Db 421 ACCTACAGCGGAGCGGGAAGACCGCGGAGGTGCTCCGAGAGGCGCACACGCTGTG 480
QY 1984 GTGCAGCTCAGACAGAGAGACCGTCTGCTGAGCGCTACAGTGACATCTGGCAGCG 2043
Db 481 GTGCAGCTCAGACAGAGAGACCGTCTGCTGAGCGCTACAGTGACATCTGGCAGCG 540
QY 2044 GCCCTGTGCACCAATCCAAACCTGATAGAGCTGTCTGTACCGAAATGCCCTGGGCGC 2103
Db 541 GCCCTGTGCACCAATCCAAACCTGATAGAGCTGTCTGTACCGAAATGCCCTGGGCGC 600
QY 2104 CGGGGGTGAAGCTGCTCTGTCAAGACTCAGACACCCCAACTGCAAACTTCAGAACCT 2162
Db 601 CGGGGGTGAAGCTGCTCTGTCAAGACTCAGACACCCCAACTGCAAACTTCAGAACCT 660
QY 2163 GAGCTGAAGAGTGGCGCATCTCCAGCTCAGCTGCGAGGACCTCTCTGCGAGCTCTCAT 2222
Db 661 GAGCTGAAGAGTGGCGCATCTCTCAGCTCAAGCTGCGAGGACCTCTCTGCGAGCTCTCAT 720
QY 2223 AGCCAA 2228
Db 721 AGCCA 726
```

```
RESULT 3
BQ900330
LOCUS
DEFINITION BQ900330 937 bp mRNA linear EST 16-AUG-2002
5', mRNA sequence.
ACCESSION BQ900330.1 GI:22292344
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 937)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2570 row: 1 column: 11
High quality sequence start: 5
High quality sequence stop: 502.
Location/Qualifiers
1. 937
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6381082"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 40"
/notes="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using zap-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
```

FEATURES

```
source
1. 937
AW673661/c
LOCUS
DEFINITION AW673661 571 bp mRNA linear EST 11-APR-2000
similar to TR:075434 O75434 ANGIOTENSIN/VASOPRESSIN RECEPTOR
AII/AVP. ; mRNA sequence.
ACCESSION AW673661
VERSION AW673661.1 GI:7538896
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 571)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Other_ESTs: ba56e01.xl
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
```

ORIGIN

Query Match 21.2%; Score 660.2; DB 5; Length 937;

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Washington University Genome Sequencing Center
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 image.llnl.gov/image/html/iresources.shtml

Possible reversed clone: similarity on wrong strand
 Seq primer: -40RP from Gibco
 High quality sequence stop: 426.

FEATURES

source
 1..571
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2900568"
 /cell_line="MGC36"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC 10"
 /notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.5 kb. Library prepared by Life
 Technologies."

ORIGIN

Query Match 17.4%; Score 539.8; DB 2; Length 571;
 Best Local Similarity 97.7%; Pred. No. 8e-121;
 Matches 558; Conservative 0; Mismatches 12; Indels 1; Gaps 1;
 QY 1297 GGGTTGTGCTCTTGGCGGAGATGGCTCTGGAATC-AGAAATCTATTTCAGGAGCA 1355
 Db 571 GGGTTGTGCTGATGCTGGCAGATGGCTCTGAAATCAATACCATCTAATGTCAGCAGCA 512
 QY 1356 GGACCTCCGGAAGCAGCGCTAGACGGGAAGAGCTCTCTGCCTTCTTCAACATGAACAT 1415
 Db 511 GGACCTCCGGAAGCAGCGCTAGACGGGAAGAGCTCTCTGCCTTCTTCAACATGAACAT 452
 QY 1416 CTTCCAGAGGACATCAACTGTGAGAGTACTACAGCTTCTACCTTGAGTTTCAGGA 1475
 Db 451 CTTCCAGAGGACATCAACTGTGAGAGTACTACAGCTTCTACCTTGAGTTTCAGGA 392
 QY 1476 ATCTTTGCAGTATGTAATATATCTTGGACGGGGAGGGGGGAGCCAGCCAGACCA 1535
 Db 391 ATCTTTGCAGTATGTAATATATCTTGGACGGGGAGGGGGGAGCCAGCCAGACCA 332
 QY 1536 GGACGTGACAGGCTGTGACCGAGTAGCGGTTTCTGAAAGAGCTTCTGCACTCAC 1595
 Db 331 GGACGTGACAGGCTGTGACCGAGTAGCGGTTTCTGAAAGAGCTTCTGCACTCAC 272
 QY 1596 CAGCGCTTCTGTTTGGACTCTTGAACGAGGAGCAGGAGCCACCTGGAGAGAGTCT 1655
 Db 271 CAGCGCTTCTGTTTGGACTCTTGAACGAGGAGCAGGAGCCACCTGGAGAGAGTCT 212
 QY 1656 CTGCTGGAAGTCTCGCGGCAATCAAGATGACCTGTTCAGTGGATCAAAAGCAAAGC 1715
 Db 211 CTGCTGGAAGTCTCGCGGCAATCAAGATGACCTGTTCAGTGGATCAAAAGCAAAGC 152
 QY 1716 TCAGAGCAGCGCTCCACCTTCGACGAGGCTCTTGGAGTCTTTCAGCTGCTGTACGA 1775
 Db 151 TCAGAGCAGCGCTCCACCTTCGACGAGGCTCTTGGAGTCTTTCAGCTGCTGTACGA 92
 QY 1776 GATCCAGGAGGAGGATTTATCCAGCAGGCCCTGAGCCACTTCCAGGTGATCGTGGTCAG 1835
 Db 91 GATCCAGGAGGAGGATTTATCCAGCAGGCCCTGAGCCACTTCCAGGTGATCGTGGTCAG 32
 QY 1836 CAACATTCCTCCAGATGGAGCAATGGTC 1866
 Db 31 CAACATTCCTCCAGATGGAGCAATGGTC 1

RESULT 5
 AG067278 664 bp DNA linear GSS 03-NOV-2001
 LOCUS

DEFINITION

ACCESSION AG067278
 VERSION AG067278.1 GI:16619080
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE

AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of Library PTB

TITLE

JOURNAL Unpublished
 (bases 1 to 664)

REFERENCE

AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission

TITLE

JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimbes@sc.riken.go.jp, URL: http://bgp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)

COMMENT

Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.

PRIMERS

Sequencing: M13Rev

LIBRARY

Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI.

FEATURES

source
 1..664
 Location/Qualifiers
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="PTB-057D11.R"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match 16.7%; Score 518.8; DB 9; Length 664;
 Best Local Similarity 92.7%; Pred. No. 1.2e-115;
 Matches 555; Conservative 0; Mismatches 43; Indels 1; Gaps 1;
 QY 750 GCTCTCCAGGAGCTCATCCGAGTTCCCGAGCGCTCTTTTCATCATCGAGGCTTCGA 809
 Db 49 GCATGCTCAGAGCTCATCCGAGTTCCCGAGCGCTCTTTTCATCATCGAGGCTTCGA 108
 QY 810 TGAGCTCAAGCCCTCTTTTCCAGGATCCTCAGGACCCCTGTCCTCTGTGGAGAGAA 869
 Db 109 TGAGCTCAAGCCCTCTTTTCCAGGATCCTCAGGACCCCTGTCCTGTGGAGAGAA 168
 QY 870 ACGCCCAAGAGCTGCTCTTTAAAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATC 929
 Db 169 ACGCCCAAGAGCTGCTCTTTAAAGCTTAATTCGGAAGAGCTGCTTTTGTGAGCTATC 228
 QY 930 TTTTGTCTCATCCACACGCGCCACCGCTTTGGAGAAGCTCCACCGCTCTGTGGAGCAACC 989
 Db 229 TTTTGTCTCATCCACACGCGCCACCGCTTTGGAGAAGCTCCACCGCTCTGTGGAGCAACC 288
 QY 990 CAGGCATGTGGAGATCCTGGGCTTCTCTGAGGAGAAAGAAATATCTTACAGTA 1049
 Db 289 CAGGCATGTGGAGATCCTGGGCTTCTCTGAGGAGAAAGAAATATCTTACAGTA 348
 QY 1050 TTTTCCCAATGTCAGAGCAGCGGGCCAGCTTTCATTTACGTGAGGAGCAACGAGCTCT 1109
 Db 349 TTTTCCCAATGTCAGAGCAGCGGGCCAGCTTTCATTTACGTGAGGAGCAACGAGCTCT 408
 QY 1110 CTTTCCCATGTGCTTCGTCCTCCCTGCTGGTGTGTGTACCTGCTCCAGCAGCA 1169
 Db 409 CTTTCCCATGTGCTTCGTCCTCCCTGCTGGTGTGTGTACCTGCTCCAGCAGCA 468

Pan troglodytes DNA, clone: PTB-057D11.R, genomic survey sequence.
 AG067278
 AG067278.1 GI:16619080

GSS.
 Pan troglodytes (chimpanzee)

Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of Library PTB

Unpublished
 2 (bases 1 to 664)

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimbes@sc.riken.go.jp, URL: http://bgp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.

PRIMERS

Sequencing: M13Rev

LIBRARY

Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI.

Location/Qualifiers

1..664
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="PTB-057D11.R"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

Query Match 16.7%; Score 518.8; DB 9; Length 664;
 Best Local Similarity 92.7%; Pred. No. 1.2e-115;
 Matches 555; Conservative 0; Mismatches 43; Indels 1; Gaps 1;
 QY 750 GCTCTCCAGGAGCTCATCCGAGTTCCCGAGCGCTCTTTTCATCATCGAGGCTTCGA 809
 Db 49 GCATGCTCAGAGCTCATCCGAGTTCCCGAGCGCTCTTTTCATCATCGAGGCTTCGA 108
 QY 810 TGAGCTCAAGCCCTCTTTTCCAGGATCCTCAGGACCCCTGTCCTCTGTGGAGAGAA 869
 Db 109 TGAGCTCAAGCCCTCTTTTCCAGGATCCTCAGGACCCCTGTCCTGTGGAGAGAA 168
 QY 870 ACGCCCAAGAGCTGCTCTTTAAAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATC 929
 Db 169 ACGCCCAAGAGCTGCTCTTTAAAGCTTAATTCGGAAGAGCTGCTTTTGTGAGCTATC 228
 QY 930 TTTTGTCTCATCCACACGCGCCACCGCTTTGGAGAAGCTCCACCGCTCTGTGGAGCAACC 989
 Db 229 TTTTGTCTCATCCACACGCGCCACCGCTTTGGAGAAGCTCCACCGCTCTGTGGAGCAACC 288
 QY 990 CAGGCATGTGGAGATCCTGGGCTTCTCTGAGGAGAAAGAAATATCTTACAGTA 1049
 Db 289 CAGGCATGTGGAGATCCTGGGCTTCTCTGAGGAGAAAGAAATATCTTACAGTA 348
 QY 1050 TTTTCCCAATGTCAGAGCAGCGGGCCAGCTTTCATTTACGTGAGGAGCAACGAGCTCT 1109
 Db 349 TTTTCCCAATGTCAGAGCAGCGGGCCAGCTTTCATTTACGTGAGGAGCAACGAGCTCT 408
 QY 1110 CTTTCCCATGTGCTTCGTCCTCCCTGCTGGTGTGTGTACCTGCTCCAGCAGCA 1169
 Db 409 CTTTCCCATGTGCTTCGTCCTCCCTGCTGGTGTGTGTACCTGCTCCAGCAGCA 468

```

Qy 1170 GCTGAGGGTGGGGGCTGTTGAGACAGACGTCAGGACCACTGCACTGCAGTGTACATGCT 1229
Db 469 GCTGAGGGTGGGGGCTGTTGAGACAGACGTCAGGACCACTGCACTGCAGTGTACATGCT 528
Qy 1230 CTACTCTGTGAGTCTGATGCAACCAAGCCGGGGG-CCCGGGCTCCAGCCCCCACCACCA 1288
Db 529 CTACTCTGTGAGTCTGATGCAACCAAGCCGGGGG-CCCGGGCTCCAGCCCCCACCACCAAGGT 588
Qy 1289 ACCAGAGAGGGTGTGCTCTTGGCGGAGATGGCTCTGGAATCAGAAAAATCTATT 1347
Db 589 TGCANAGAGGGTGTGCTCTTGGCGGAGATGGATCGAATCAGAAAAATCTATT 647

RESULT 6
LOCUS AW956628 568 bp mRNA linear EST 01-JUN-2000
DEFINITION EST368698 MAGE resequencings, MAGD Homo sapiens cDNA, mRNA sequence.
ACCESSION AW956628
VERSION AW956628.1 GI:8146311
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (Bases 1 to 568)
Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspar,R., Gay,C.,
Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
Quackenbush,J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: john@tigr.org
Plate: 94
Seq primer: Reverse.
FEATURES
source Location/Qualifiers
1..568
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequencings, MAGD"
/note="Vector: pbluescriptSKm"
ORIGIN
Query Match 15.9%; Score 494.2; DB 2; Length 568;
Best Local Similarity 97.4%; Pred. No. 1.2e-109;
Matches 523; Conservative 0; Mismatches 10; Indels 4; Gaps 2;
Qy 1638 CCACCTGGAGAGAGTCTCTGCTGGAGGTCTCGCGGCACATCAAGATGACCTGTGCA 1697
Db 1 CCACCTGGAGAGAGTCTCTGCTGGAGGTCTCGCGGCACATCAAGATGACCTGTGCA 60
Qy 1698 GTGATCCAAAGCAAGCTCAGAGCGAGCGGTCCACCTGCAGCAGGGCTCTTTGGAGTT 1757
Db 61 GTGATCCAAAGCAAGCTCAGAGCGAGCGGTCCACCTGCAGCAGGGCTCTTTGGAGTT 120
Qy 1758 CTTCAGCTGCTTGTACGAGATCCAGGAGGAGGAGTTTATCCAGCAGGCGCTTGAGCCACTT 1817
Db 121 CTTCAGCTGCTTGTACGAGATCCAGGAGGAGGAGTTTATCCAGCAGGCGCTTGAGCCACTT 180
Qy 1818 CCAGTGTATCTGGTTCAGCAATGCTTCCAAAGATGAGACATGCTCTCTGTTCTG 1877
Db 181 CCAGTGTATCTGGTTCAGCAATGCTTCCAAAGATGAGACATGCTCTCTGTTCTG 240
Qy 1878 TCTGAAGCGCTGAGGAGCGCCAGGTGCTGACCTTGTATGCGCGCACCTACAGCGGGA 1937
Db 241 TCTGAAGCGCTGAGGAGCGCCAGGTGCTGACCTTGTATGCGCGCACCTACAGCGGGA 300

```

```

Qy 1938 CGGGGAAGACCGCGCAGGTGCTCCGAGAGAGCGACACGCTGTGTGTCAGCTCAGACC 1997
Db 301 CGGGGAAGACCGCGCAGGTGCTCCGAGAGAGCGACACGCTGTGTGTCAGCTCAGACC 357
Qy 1998 AGAGAGACCGTCTGCTGGACGCTACAGTGAACATCTGCAGCGGCGCTGTGCACCAA 2057
Db 358 AGAGAGACCGTCTGCTGGACGCTACAGTGAACATCTGCAGCGGCGCTGTGCACCAA 417
Qy 2058 TCCAAACCTGATAGAGCTGCTCTGTACCCGAAATGCGCTGGCGAGCGGGGGTGAAGCT 2117
Db 418 TCCAAACCTGATAGAGCTGCTCTGTACCCGAAATGCGCTGGCGAGCGGGGGTGAAGCT 477
Qy 2118 GCTCTGTCAAGGACTCAGACACCCCACTGCAAACTTCAGAACTCAGAGCTGAAGAG 2174
Db 478 GCTCTGTCAAGGACTCAGACACCCCACTGCAAACTTCAGAACTCAGAGCTGAAGAG 533

RESULT 7
LOCUS CV030364 669 bp mRNA linear EST 20-AUG-2004
DEFINITION 9461 Full Length cDNA from the Mammalian Gene Collection Homo
sapiens cDNA 5' similar to BC028069, mRNA sequence.
ACCESSION CV030364
VERSION CV030364
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (Bases 1 to 669)
Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
Dricot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,
Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,
Simmons,B., Sequerre,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C.,
Vandenhaute,J., Cusick,M.E., Albaladejo,J.S., Hill,D.E. and Vidal,M.
Human ORFeome Version 1.1: a Platform for Reverse Proteomics
Genome Res. (2004) In press
Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
results from a PCR reaction using an MGC full-length cDNA as
template DNA and ORF specific primers
PCR Primers
FORWARD: ATGCTACGAACCCGAGGACG
BACKWARD: CAGCAGCCAATGTCCAAA
Insert Length: 669 Std Error: 69.00
Plate: 11079 row: 09 column: F
Seq primer: ACTGGCGCTGTTTACACAGCTCGTGTGCTGGGAAAC
High quality sequence start: 98
High quality sequence stop: 668
POLYA=No.
FEATURES
source Location/Qualifiers
1..669
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="mixed"
/clone_lib="Full Length cDNA from the Mammalian Gene
Collection"
/note="vector: mixed; The ORFs were PCR amplified from the
MGC (Mammalian Gene Collection) as of April 2004 and
cloned by recombinational Gateway cloning into pDONR223
Donor vector. Reference : MGC (Mammalian Gene Collection)
Program Team, Generation and Initial Analysis of more than
15,000 Full-length Human and Mouse cDNA Sequences. PNAS,
2002, 99(26), 16899-16903"
ORIGIN

```


REFERENCE
AUTHORS
CONSRW
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 4123)
Koehler, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Oeinger, A.,
Fobo, G., Han, M. and Wiemann, S.
The German cDNA Consortium
Direct Submission
Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the
Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA
sequencing consortium of the German Genome Project. This clone
(DKFZ58601822) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZ58601822
Further information about the clone and the sequencing project is
available at <http://mips.gsf.de/projects/cdna/>.

FEATURES
source

1. 4123
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZ58601822"
/db_xref="taxon:9606"
/clone="DKFZ58601822"
/tissue_type="uterus"
/clone_lib="586 (synonym: hutei). Vector pSport1; host
DH10B; sites NotI + SalI/MluI"
/dev_stages="adult"
/note="death of cell filament-forming Ced-4-like
apoptosis protein, N-terminus truncated"

gene
CDS

1. 4123
/gene="DKFZp58601822"
2. 3580
/gene="DKFZp58601822"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB55945.1"
/db_xref="GI:5911939"
/db_xref="GOA:O9C000"
/db_xref="UniProt/Swiss-Prot:O9C000"
/translation="LLQLRPHRPSQDPLVKKSPDYPVEENRHHLEIRDLFGPLDQ
EPRIIVLQAGIKSTLQARKEAGQGLYGRFQHFYFSCRELAQKWSLAL
IGKGTATPAPRIQLSRPRLILFDGVDPEPVLQEPSELCLHWSQPQADALIG
SLGKTLPEASPLITARTALONLIPLEOARWVVLGFSERKEFYFYRFTDQ
AIRAPLVKSNKELWALCLVPWVSLACTCLMOOKRKEKLTLSKTTTLCLHYLAQ
ALQAPLQPLRLCSLAEGIWQKTLFSPDDLRKGLDGLAISTFLKMLQEHFI
PLSYFHLPCQEFPAAMSVYLEDKGRGHSNCTIDLEKTLAYGHLGLFASRTF
LGLSLDEGEREMENI FHCRLQGRNLMQWVPSLQLLQPHSLSLCLYETRNKTL
TQNAHFEEMCMVETDMLLVCTFCIKFSRHHVKLLQLEGRQHRSTWSPMTVLRFL
VPVTDWQILFVSLVTRNLKELDLSGNLSHSAVSKLCTLRPRCLLETLLRAGC
GLTAECDKDLAFLGRANQTLTDLSENVLTDAKHLQRLRQPSCKLQRLQVSG
LTSQCDLAVLSASLSKELDLQNNLDVGRLLCEGLRHACKLRLGLDQITL
SDENQRLEALDEQKOLLIFSRKSPVMTPEGLDGMNSNTSLKRLQRLGSRRA
SHVAQNLKLDLVSKIPIFAIESSEPPVVELLCVPSASQGLDHTPLGTDGDF
WGPTGVATEVVDKKNLYVHFVAGVSPWNTGLCFVNRVAVTVFVFCVWDQFQ
EINPQHSWVAGPLDITKABEPAVEAVHLPHFVQLQGHVDTSIFQVAHFEKEMLE
KPARVELHIVLENPSPGLVLLKMLHNAIRLPIPTSVLLVHRVPEVTHYLI
PSCSIRKALDLENKTFQVRIHKLPPLETLDMGCRVYSGSGSMELIIPKLELCY
RSPGQDLQFSEFYHGLSGIRLQVKDDLETLVMEALVPGDMPATTIIPARIAY
PSPLDAPQLHFDVQYREQLIARTVTSVEVLDKLHGVLSQEQYVERVLAENTRPSQMR
KLFLSLQSWDRKCKDGLYQALKEHPLHLMELWEKSKGLLPLSS"

ORIGIN

Query Match 12.8%; Score 398; DB 3; Length 4123;
Best Local Similarity 53.6%; Pred. No. 7.5e-86;
Matches 1082; Conservative 0; Mismatches 85; Indels 81; Gaps 9;
Qy 539 CCGAGCCACCGCGACCTGCTCATGCAAGCGCGGCGGATAGGCAAGTCCATGCTGG 598
Db 129 CCCAAGAACCTCGCATAGTCATGCGAGGGGGCTGCTGGAAITGGGAAGTCAACACTGG 189

Qy 599 CACACAAGGTGATGCTCGATCTGGCGCGAAGCTCTTCCAAAGGAGATTTGATATC 658
Db 189 CCAGCGAGGTGAAGAAACCTTGGGGGAGAGCCAGCTGTATGGGAGCCGCTTCCAGCATG 248
Qy 659 TCTTCTACATCAACTGCAGGAGATGAACAGAGTGCACGGAATGAGCATGCAAGACC 718
Db 249 TCTTCTACTTTCAGCTGCAGAGAGCTGCCAG---TCCAAGGTGGTGAATGCTCGCTGAGC 305
Qy 719 TCATCTTTCAGCTGCTGGCCTGAGCCAGCGCCTCTCCAGGAGCTCATCCGAGTTCCTCG 778
Db 306 TCATCGGAAGAAGATGGGACACCACTCCGGCTCCATTAGACAGATCTGTCTAGGCCAG 365
Qy 779 AGCCCTCTCTTTCATCATCGAGCTTCGATGAGCTCAAGCTCTTCTTCCACATCTCTC 838
Db 366 AGCGCTGCTCTTTCATCTCGATGCTGTAGATGAGCCAGGATGGGTCTTTCAGGAGCGGA 425
Qy 839 AGGGACCTTGGTGCCTCTGCTGGGAGGAGAAACGCGCCACCGAGCTCTCTTCAACAGCT 898
Db 426 GTTCTGAGCTCTGTCTGCATCTGGAGCCAGCCAGCGCGGATGCACTGCTGGGCGATT 485
Qy 899 TAATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGCGCCACCGCTT 958
Db 486 TGCTGGGGAACCTATATCTTCCGAGGATCTTCTCTGATCACGGCTCGGACACAGCTC 545
Qy 959 TGGAGAAGTCCACCGTCTGCTGGAGCACCCAGGAGATGAGAGATCTCTGGGCTTCTCTG 1018
Db 546 TGCAGAACCTCATTTCTTCTTGGAGCAGGACGCTGGGTAGAGGTCTCTGGGGTCTCTG 605
Qy 1019 AGGCAGAAAGGAGGATCTTCTCAAGATATTTCCAAATGACAGAGCGGGGGCCCAAG 1078
Db 606 AGTCCAGAGGAGGAGGATATTTCTACAGATATTTTACAGATGAAAGCAAGCAATTAGAG 665
Qy 1079 TCTTCAATACGTGAGGAGCAACGAGCTCTCTTACCATGTCTTCCCTCCCTCGTGT 1138
Db 666 CTTTAGTGTGGTCAATCAACAAAGAGCTCTGGGCTCTGTGTCTTGTGCTCTGGGTGT 725
Qy 1139 GTGGGTGTGTGTACTGCTCCAGCAGCTGGAGGCTGGGGGCTGTGTGAGACAGA 1198
Db 726 CTTGGCTGGCTGCTGCTGCTGATGAGCAGATGAAGCGAGGAGGAGGAGGAGGAGGAG 785
Qy 1199 CGTCCAGGACCACTGAGTGTACATGCTCTACCTGCTGAGTCTCATGCAACCAAGC 1258
Db 786 CTTTCAAGACCAACCAACCTCTGTCTATCTTACCTTGGCCAGGCTCTTCCAAGCTCAGC 845
Qy 1259 CGGGGGCCCGCGCTCCAGCCCAACCAAGAGAGGCTGTGCTCTTGGCGGCGAG 1318
Db 846 CATTTG-----GACCCAGCTCAGAGACCTCTGCTCTCTGCTGCTGTG 887
Qy 1319 ATGGCTCTGGAATCAGAAAATCTATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1378
Db 888 AGGCGATCTGGCAAAAAGAGACCTTTTCAAGTCCAGATGACCTCAGGAGGAGGAGGAGGAG 947
Qy 1379 ACGGGGAGAGGCTCTGCTGCTTCTCAACATGAACATCTTCCAGAGGAGGAGGAGGAGGAG 1438
Db 948 ATGGGGCCATCATCTCCACCTTTTGAAGATGGGTATTTCTTCAAGAGGAGGAGGAGGAGGAG 1007
Qy 1439 AGAGGTACTACAGCTTCTCATCTGAGTTTTCAGGAATTTCTTTCAGGCTATGTACTATA 1498
Db 1008 TGAG---CTACAGCTTCATTCACCTCTGTTTCCAGAGGTTCTTTCAGGAGGAGGAGGAGGAG 1064
Qy 1499 TCTTGGACGAGGGGAGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1558
Db 1065 TCTTGGAGGATGAGAGGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1124
Qy 1559 AGTACGGTTTTCTGAAGGAGGCTTCT-----GGCACTCAACAGCGCTTCTCTGTTG 1612
Db 1125 CGTTAGAGGATATGGAATATCATGGCTGTGTGGGGCATCAACCCACACGCTTCTCTATTTGG 1184
Qy 1613 GACTCTTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1672
Db 1185 GCCTGTTAAGTGTAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1244
Qy 1673 CGCACATCAAGATGGACCTGTTGAGTGGATCCAAAGCAAGCTCAGAGGAGGAGGAGGAGGAGGAG 1732

Db 1245 AGGG-----GAGAACCTGATGAGTGGTCCGTCCTCGAGCTGCT----- 1287
Qy 1733 CCTGACGAGGCTCTCTTGGAGTCTTTCAGTCTGTTGACGAGATCCAGGAGGAGAT 1792
Db 1288 -GCTGACGACACTCTCTGAGTCCCTCCTCACTGCTTGTACGAGACTCGGAACAAACCT 1346
Qy 1793 TTATCCAGAGGCGCTGAGCACTTCAGGTGATCGTGTGACGACCAATTCGCTCCAAGA 1852
Db 1347 TCCTGACACAAGTGAATGGCCCATTTGGAAGAAA---TGGGCATGTGTGTAGAAACAGACA 1403
Qy 1853 TGGAGCATATGCTCTCTCGTCTGCTGAAGCGCTGCGAGAGCGGCCAGGTCGTCACT 1912
Db 1404 TGGAGCTCTTAGTGTGACCTTTCTGCATTAATTCAGCGCCACGCTGAGAACTTCAGC 1463
Qy 1913 TGTATGCGGCCACTACAGCGGAGCGGGGAAGACCGCGGAGGTCTCGCAGGAGCGC 1972
Db 1464 TGATT-----GAGGCGAGGCACACAGATCAACATGAGGAC 1499
Qy 1973 ACAGCTGTTGGTCAGCTCAGACGAGAGGACCGTTCTGCTGGAGCGCTACAGTGAAC 2032
Db 1500 CCACCATGGTAGTCTCTTTCAG-----GTGGGTCCCGAGTCACAGATGCCCTATTGGCAGA 1553
Qy 2033 ATCTGGCAGCGGCTGTGCAACCAATCCAAACCTGATAGAGCTGTCTCTGTACCGAAATG 2092
Db 1554 TTCTCTTCTCGTCTCAAGGTCACACAGAAACCTGAAGGAGCTGGACCTAAGTGGAACT 1613
Qy 2093 CCCTGGGAGCGCGGGGGTGAAGTCTCTGTCAAGGACTCAGACACCCCACTGCAAC 2152
Db 1614 CGCTGAGCCACTCTGCACTGAAGAGTCTTTGTGAAGACCTTGTGAAGACCTCGCTGCTCC 1673
Qy 2153 TTCAGAACCTGAGGCTGAAGAGTGGCGCATCTCCAGCTCAGCTGCGGAGGACTCTCTG 2212
Db 1674 TGGAGACCTGCGGTGGCTGTGGCTGTGGCTTCAAGCTGAGGACTGCAAGGACTTGGCT 1733
Qy 2213 CAGCTCTCATGACCAATAAGAAATTTGACAAGGATGATCTCAGTGGCAACCGGCTTGGAT 2272
Db 1734 TTGGGCTGAGACCAACAGACCCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTACGG 1793
Qy 2273 TCCAGGATGATGCTGTTTGCAGGGCTGCGGCTCCCGACTCCCGAGCTGAGGCTGAGATG 2332
Db 1794 ATGCTGGAGCCAAACACATTTGGCAGAGACTGAGACAGCGGAGCTGCAAGCTACAGGAC 1853
Qy 2333 TTCAGTTGAGGAAGTGTGAGCTGAGTTCGGGGCTTGTGAGGAGATGGCTTCTGTGCTG 2392
Db 1854 TGCAGCTGCTGAGTGTGGCTCAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1913
Qy 2393 GCACCAACCCACATCTGTTGAGTGGACCTGACAGGAATGCACTGGAGATTTGGGCC 2452
Db 1914 GTGCCAGCCCGAGCTGAAGGAGCTAGACCTGCAGCAGAACCACTGGATGACGCTTGGCG 1973
Qy 2453 TGAGTTACTATGCGAGGACTGAGGACCCAGCTTCGAGACTACGAGCTTTGTGGCTGA 2512
Db 1974 TGCAGCTGCTGTGAGGGGCTCAGGCACTCTGCTGCTGCAAACTCATAGCCTGGGGCTGG 2033
Qy 2513 AGATCTGCGGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2550
Db 2034 ACCAGACACTGTGATGATGATGAGGAGGAGGAGT 2071

RESULT 13
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
A0094928
HS_3028_A2_A07 MF CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=3028 Col=14 Row=A, genomic survey
sequence.
A0094928
A0094928.1 GI:3466375
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
1 (bases 1 to 396)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3028 row: A column: 14
Class: BAC ends
High quality sequence stop: 396.
Location/Qualifiers
1..396
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=3028 Col=14 Row=A"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

FEATURES
source
1..396
Query Match 12.7%; Score 395; DB 8; Length 396;
Best Local Similarity 99.7%; Pred. No. 2.1e-85;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1541 TGACACAGGCTGTGACCGAGTACGCGTTCCTGAAGGAGCTTCCTGGCCTCACCAGCC 1600
Db 1 TGACACAGGCTGTGACCGAGTACGCGTTCCTGAAGGAGCTTCCTGGCCTCACCAGCC 60
Qy 1601 GCTTCTCTGTTTGGACTTCCTGAACGAGGAGACGAGCCACCTGGAGAGAGTCTCTGCT 1660
Db 61 GCTTCTCTGTTTGGACTTCCTGAACGAGGAGACGAGCCACCTGGAGAGAGTCTCTGCT 120
Qy 1661 GGAAGTCTCGCGGACATCAAGATGGACCTGTTGAGTGGATCCAAAGCAAAGTCTCAGA 1720
Db 121 GGAAGTCTCGCGGACATCAAGATGGACCTGTTGCACTGGATCCAAAGCAAAGTCTCAGA 180
Qy 1721 GCGACGGCTCCACCTTCGACGAGGCTCCTTCAGTTCCTTCAGTCTGTTGACGATCC 1780
Db 181 GCGACGGCTCCACCTTCGACGAGGCTCCTTCGAGTTCCTTCAGTCTGTTGACGATCC 240
Qy 1781 AGGAGGAGGAGTTTATCCAGAGGCGCTGAGCCACTTCAGGTTGATCGTGGTCAGCAACA 1840
Db 241 AGGAGGAGGAGTTTATCCAGAGGCGCTGAGCCACTTCAGGTTGATCGTGGTCAGCAACA 300
Qy 1841 TTGCTCTCAAGATGGAAGCACTGCTCTCTGTTCTGTTCTGAGCGCTCAGGAGCGCCC 1900
Db 301 TTGCTCTCAAGATGGAAGCACTGCTCTCTGTTCTGTTCTGAGCGCTCAGAGCGCCC 360
Qy 1901 AGTGTCTGCACTTGTATGGCGCCACCTACAGCGCGG 1936
Db 361 AGTGTCTGCACTTGTATGGCGCCACCTACAGCGCGG 396
RESULT 14
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
BC021272
Homo sapiens cDNA clone IMAGE:4811303, containing frame-shift
errors.
BC021272
BC021272.2 GI:33878145
HTC.

| | |
|-----------|--|
| Source | Homo sapiens (human) |
| Organism | Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| Reference | 1 (bases 1 to 3359) |
| Authors | Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klauener,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altshuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Uesdin,T.B., Toshiyuki,S., Carninci,P., Frange,C., Raha,S.S., Joquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A. |
| Title | Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences |
| Journal | Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) |
| Pubmed | 12477932 |
| Reference | 2 (bases 1 to 3359) |
| Authors | Director MGC Project. |
| Title | Direct Submission |
| Journal | Submitted (14-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA |
| Remark | NIH-MGC Project URL: http://mgc.nci.nih.gov |
| Comment | On Aug 19, 2003 this sequence version replaced gi:18204229. Contact: MGC help desk Email: cgabbs@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC) Gaithersburg, Maryland; Web site: http://www.nisec.nih.gov/ Contact: nisc_mgc@hgr.nih.gov Akhter,N., Avale,K. Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,A., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D. |
| Features | Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Project: 39 Row: 9 Column: 3 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19745161 This clone has the following problem: frame shifted. Location/Qualifiers 1..3359 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4811303" /issue_type="Placenta" /clone_lib="NIH MGC 21" /lab_host="DH10B-R" /note="Vector: pOTF7" |
| Origin | |

| | | | |
|----|------|--|------|
| Db | 1486 | AGCTGTGAGATGTTACAGGAAATGCTAGTGCCTAATTTTGAAGAACAGAGGAGCA | 1545 |
| Qy | 1584 | CTGGCACTCACAGCGCTCTCTGTTGGACTCTCTGAACAGGAGACACAGGAGCCACT | 1643 |
| Db | 1546 | TTGATTTTTTTGGGGTGTCTTCTAAGTGGCTTTTAAATATAAAGAAACAAAGAACT | 1605 |
| Qy | 1644 | GGAGAGAGTCTCTGCTGGAAGTCTGCGGCAATCAAGATGAGACTGTGAGTGGAT | 1703 |
| Db | 1606 | GGATGCGTTTTTGGCTTCCAACTGTCCCAAGAGATAAAGCAGCAAAATTCACAGTGCCT | 1665 |
| Qy | 1704 | CCAAAGCAAGCTCAGAGCGAGCGCTCCACCTGCGAGCGAGGCTCTCTGAGTTCCTCAG | 1763 |
| Db | 1666 | GAAAGAGTTTGGGAGCGTGGCAATCTCAGGAGCAGTGGATTCCTTGGCGATATTTA | 1725 |
| Qy | 1764 | CTGCTTGTACAGATCCAGGAGGAGGAGTTTATCCAGAGCGCTCAGGCACTTCAGGT | 1823 |
| Db | 1726 | CTGTCTCTTGAATGAGGATCTGCTGCTTTGTGAAGCAGGAGTGAACCTCTCCAAAG | 1784 |
| Qy | 1824 | GATCGTGTGAGCAACATTTGCTCTCAAGATGAGACATGCTCTCTGCTCTGCTGAA | 1883 |
| Db | 1785 | --AAGCTAACTTTCATATATTATTAACAACGTGGACTTGGTGGTTTCTGCGCTACTGCTTAA | 1842 |
| Qy | 1884 | CGCTGCGAGGAGCGCCAGGTGCTGCACTTGTATGGCGCCACCTACAGCGCGACGCGGA | 1943 |
| Db | 1843 | ATACTGCTCCAGCTTGAGGAACTCTGTTTTTCCGTTCAAAATGCTTTAAGAAAGAGA | 1902 |
| Qy | 1944 | AGACCGCGAGAGTCTCCGAGAGCGCACACGCTGTTGGTGAGCTCAGACAGAGAG | 2003 |
| Db | 1903 | TGACACAGCTTACGTCGGATTACAGCTCATCTGTTGGCATCACATCTGCTCTGTG-- | 1960 |
| Qy | 2004 | GACCGTCTGTGAGCGCTACAGTGAACATCTGGCAGCGGCTGTGACACCAATCCAA | 2063 |
| Db | 1961 | -----CTCACACACGCGGGCA | 1977 |
| Qy | 2064 | CCTGATAGCTGTCTCTGTATCCGAAATGCCCTGGCAGCGGGGGTGAAGCTGCTG | 2123 |
| Db | 1978 | CCTCAGAGCTCCAGGTGAGGACAGCACCTCTCAGCGAGTGCACCTTTGTGACCTGGT | 2037 |
| Qy | 2124 | TCAGGACTCAGACACCCCACTGCAAACTTCAGAACTGAGGCTGAAGAGTGGCCCAT | 2183 |
| Db | 2038 | TAAACAGCTGAGGATCCAGCTGTGCGCTTCAGAACTTGAATAAATAACGTTTCTCT | 2097 |
| Qy | 2184 | CTCAGCTCAGCTCGAGGAGCTCTCTGAGCTCTCATGCCAATGAATTTGACAAG | 2243 |
| Db | 2098 | TTCTGGCCAG---AGTGTCTGCTTTGAGGTGCTCTTTATCAGCCAGCTTGAATA | 2154 |
| Qy | 2244 | GATGATCTCAGTGGCAACGGCTTGGATTCCAGGCAATGATGCTTTTCGAGGCGCT | 2303 |
| Db | 2156 | CCTGAGCTTCACTTACGMAACTCTCTCGTGATGACATCAGGTCCCTCTGTGATGCTT | 2214 |
| Qy | 2304 | GGGGATCCCAAGTGCAGGCTGCAGATGATTCAGTTGAGGAGTGTGAGTGGAGTCCGG | 2363 |
| Db | 2215 | GAACTACCCAGCAGGCAACGTCAGAGAGTACGCTGCTGTAATTTGTACCTCTCACCAT | 2274 |
| Qy | 2364 | GCCTGTGAGGAGTGGCTTCTGCTGCTGGCAACCAACCAATCTGTTGAGTTGAGCT | 2423 |
| Db | 2275 | TGATTTGAGTCTCTGCTGGCTTCTAACCAACCAAGAGCTGACGTATCTGAATGT | 2334 |
| Qy | 2424 | GACAGAAATGCACTGGAGGATTTGGGCTGAGGTTATATGCGAGGAGCTGAGCAGCC | 2483 |
| Db | 2335 | ATCCTGCAA---CCAGTTAGACACAGGCGTGGCCCTTTTGTGTGAAGCCCTGTGAGGCC | 2391 |
| Qy | 2484 | AGTCTGCAGACTACGACTTTGCTGCTGAAGATCTCGCGCTCTCACTGCTGCTGTGTA | 2543 |
| Db | 2392 | AGACAGGCTCTGGTATACCTGATGTGGCTTCTGCGACCTTCCAGGAGCTGTGCGA | 2451 |
| Qy | 2544 | CGAGCTGGGCTCAACTCTCAGTGTGAACCAAGAGCTGTGAGAGCTGTGAGCTGAA | 2603 |
| Db | 2452 | ATACATCTCTGAAATGCTTCTGCGTAAACAAGAGGCTGCGCTATCTAGACCTCAGTGCAA | 2511 |
| Qy | 2604 | TGAGCTGGGGACCTCGGGGTGCTGCTGTGTGAGGCGCTCAGGCTCCACGTGCAA | 2663 |

| | | | |
|----|------|--|------|
| Db | 2512 | TGCTCTGAAGGACGAGGACTGAAAACTCTCTGCGAGGCTTGAAACATCCGAGCTCTG | 2571 |
| Qy | 2664 | GCTCCAGACCTCGGTTGGCATCTCGCGCTGGGCTCTCGCGCTGTGAGGCTCTTTC | 2723 |
| Db | 2572 | CCTGGATTCATCTGTTGGTAAATGTTTATCACTGCTGCTGCTGTGAGACCTCGC | 2631 |
| Qy | 2724 | TGTGTGCTCCAGGCGCAACCACTCCGCGAGCTTGAGTTTCAACGACCTGGG | 2783 |
| Db | 2632 | CTCTGCTCTCATCAGCAATCAAACTTGAAGATTCTGCAAAATGGTGCAATGAATCGG | 2691 |
| Qy | 2784 | AGACTGGGCGCTGTTGCTGGCTGAGGCGCTGCAACATCCCGCTGCGAGCTCCAGAA | 2843 |
| Db | 2692 | AGATGCGGCTGTCAGCTGTTGTGTCGGGCTCTGACGCACTACGATTGCCCTTAGAGAT | 2751 |
| Qy | 2844 | ACTGTGGCTGATAGCTGTCGCGCTCAGACGCAAGGCTTGTGAGAACTCTTACTTCA | 2903 |
| Db | 2752 | TCCTGGGTTGGAAGATGTTGGTTAAAGAGCACTGCTGTAAAGATCTCGGCTGTCT | 2811 |
| Qy | 2904 | GGGATCAACAGACCTTTGACCGACTTACCTGACCAACCAACGCTTGGGGACACAGG | 2963 |
| Db | 2812 | CACCTGCAGTAAGACCTTGCAGCAGCTCAACCTGACCTTGAACACCTTTGGACCA | 2871 |
| Qy | 2964 | TGTCGAGCTGTTTCCAGCGGCTGAGCATCTCGGCTGCAACTCCGAGTCTCTGTT | 3023 |
| Db | 2872 | GGTGTGTTGACTCTGTGAGGCGCTGAGACACCCAGAGTGTGCCCTGCGAGTGTCTGGGCT | 2931 |
| Qy | 3024 | ATTTGGGATGACCTGAAATAAATGACCCACAGTAGGTTGGCAGCGCTTCAGATTAACAA | 3083 |
| Db | 2932 | GAGAAACTGATTTTGTAGGAAACCCAGGCACTTCTGACGCTGAGGAAGAGAGAA | 2991 |
| Qy | 3084 | ACCTTATTGAGAT 3098 | |
| Db | 2992 | TCCTAACCTGACAT 3096 | |

RESULT 15

BC012789

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

3360 bp mRNA linear HTC 12-OCT-2004

BC012789 Homo sapiens cDNA clone IMAGE:3957172, containing frame-shift errors.

BC012789 Homo sapiens (human)

BC012789.2 GI:33872518

HTC.

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3360)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shevchenko, Y., Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D., Bouffard, G.G., Blakesley, R.W., Grimwood, J., Schmutz, J., Myers, R.M., Dickson, M.C., Rodriguez, A.C., Krzywinski, M.I., Skalska, U., Smal, D.E., Butterfield, Y.S., Krzywinski, M.I., Jones, S.J. and Marra, M.A.

Schmerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 3360)

Director MGC Project.

Direct Submission

JOURNAL

Submitted (15-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:15215377.

Contact: MGC help desk

Contact: Рос патр депт
Email: cgabbas-r@mail.nih.gov

Email: csarbbg@mail.nih.gov
Tissue procurement: ATCC

Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory

cdNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbiology.org>

contact: amadan@systemsbiology.org

Contact: anuradhasystemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha

Amur Madan, Jessica Falley, Etili Helton, Mark Kettelman, Anurfa Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 25 Row: k Column: 23
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19745161
 This clone has the following problem: frame shifted.

FEATURES

source

```
1. 3580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3957172"
/tissue_type="Placenta"
/clone_lib="NIH MGC_21"
/lab_host="DH10B-R"
note="Vector: pOTB7"
```

ORIGIN

Query Match 12.1%; Score 377; DB 3; Length 3360;
Best Local Similarity 48.9%; Pred. No. 9.9e-81;
Matches 1269; Conservative 0; Mismatches 1260; Indels 6

| | | | | | |
|-----|----|-------------------------|---------------------|------------------------------|------|
| 513 | QY | CCTCTTTGAGCCAGA | CGAGAGCGCCCGGAGCCAC | CGCGCACCGTGTGTATGCAAGCGC | 572 |
| 469 | Db | CCGCTTTTGTCTCCCAAGAAAT | CTGGAAACAGCCACG | TACAGTATTTCAAGGACC | 528 |
| 573 | QY | GGCAGGATAGGCAAGTCCATGT | TGGCACACAAGGTGATCT | CGACTGGCGGACGGAA | 632 |
| 529 | Db | ACAAGAAATTGGAAAAACGACAT | CTGTATGAAGCTGATG | CGCTTGGTCGCAACAACAA | 588 |
| 633 | QY | GCTCTTCCAAGGCAGATTGTA | TATATCTTCTTACATCAA | CTGCAGGAGATGAACACAG | 692 |
| 589 | Db | GATCTTTCGGATAGGTTCCGT | TACAGTTCTATTTCTG | CTGCAGAGAACTGAGGAG-- | 646 |
| 693 | QY | TGCCACGGAAATGACAGAT | TGCAAGACCTCATCTT | CAGCTGTGGCTGAGCCACAGCGGCC | 752 |
| 647 | Db | -TTGCGGCCCAACAGATTT | TGGCTGACTTGTATTT | CCAGAGATGGCGCTGACCCGCTGTCTCC | 705 |
| 753 | QY | TCTCCAGGAGCTCATCCGAG | TTCCGAGCGCTCTTTT | TTCATCATCGAGCGCTTCGATGA | 812 |
| 706 | Db | TATAACAGAGATCGTGCT | CTCAACCGAGAGACTCT | TGTTGTCATCGACAGCTTCGAAGA | 765 |
| 813 | QY | GCTCAAGCCCTTTTCCACGAT | CCTCTCAGGACCCCTG | TGCTCTGCTGGAGGAGAAACG | 872 |
| 766 | Db | GCTGCAGGGCGGCTTGAA | CGAACCCGATTCGGAT | CTGTGTGGTGACTTGATGAGAAACG | 825 |
| 873 | QY | GCCCACGGAGCTGCTTCTT | TAACTTAATTCGGAAGAC | TGCTGCTCGCTAGCTATCTTT | 932 |
| 826 | Db | GCCGGTGCAAGTGCCTT | CTGACAGTTTGTCTG | AGGAAGATGCTCCCGAGAGCCTCCCT | 885 |
| 933 | QY | GCTCATCACCAACACGCG | CCACGGCTTTGGAGAAGCT | CCACCGTCTGCTGGAGCACCCACG | 992 |
| 886 | Db | GCTCATGCGCATCAAA | CCCGTGTGCCGAAGGAGCT | CCGGGATCAGGTGACGATCTCAGA | 945 |
| 993 | QY | GCATGTGGAGATCCTGGG | CTTCTCTGAGGCAGAAAG | GAATACTTCTTCAAGTATTTT | 1052 |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 14:41:04 ; Search time 173 Seconds
(without alignments)
2313.855 Million cell updates/sec

Title: US-10-781-294-24
Perfect score: 5472
Sequence: 1 MLRTAGDGLCRSLTYLEEL.....MTHSLAALRVTPYLDIGC 1035

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|------------|--------------------|
| 1 | 5472 | 100.0 | 1035 | 5 ADE36417 | Adm36417 Human PAA |
| 2 | 5472 | 100.0 | 1035 | 8 ADJ19339 | Adj19339 Human PAN |
| 3 | 5448.5 | 99.6 | 1062 | 8 ADP47724 | Adp47724 Human Mon |
| 4 | 5433 | 99.3 | 1061 | 5 AAO15590 | AAo15590 Human PYR |
| 5 | 5433 | 99.3 | 1061 | 7 ADF94755 | Adf94755 Human PYR |
| 6 | 5389 | 98.5 | 1099 | 5 AAO17857 | AAo17857 Pyrin dom |
| 7 | 5259.5 | 96.1 | 1027 | 8 ADJ19383 | Adj19383 Human PAN |
| 8 | 5113.5 | 93.4 | 1006 | 8 ADP47726 | Adp47726 Human Mon |
| 9 | 5077.5 | 92.8 | 1004 | 7 ADM04476 | Adm04476 Human pro |
| 10 | 4799.5 | 87.7 | 950 | 8 ADP47728 | Adp47728 Human Mon |
| 11 | 4520 | 82.6 | 865 | 7 ADC31287 | Adc31287 Human nov |
| 12 | 4458.5 | 81.5 | 892 | 8 ADP47730 | Adp47730 Human Mon |
| 13 | 3996 | 73.0 | 1033 | 8 ADP47732 | Adp47732 Murine Mo |
| 14 | 3364 | 61.5 | 635 | 8 ADJ19399 | Adj19399 Human PAN |
| 15 | 3163 | 57.8 | 603 | 5 ADE36457 | Adc36457 Human PAN |
| 16 | 3163 | 57.8 | 603 | 8 ADJ19379 | Adj19379 Human PAN |
| 17 | 3005 | 54.9 | 582 | 6 ABU99119 | Abu99119 Novel hum |
| 18 | 3005 | 54.9 | 582 | 8 ADM93803 | Adm93803 Human nov |
| 19 | 2953.5 | 54.0 | 565 | 5 ABG97475 | Abg97475 Human nuc |
| 20 | 2694.5 | 49.2 | 521 | 6 ABU99120 | Abu99120 Novel hum |
| 21 | 2671.5 | 48.8 | 521 | 8 ADM93805 | Adm93805 Human NOV |
| 22 | 2513 | 45.9 | 472 | 8 ADP47884 | Adp47884 Human CAT |
| 23 | 2396 | 43.8 | 1034 | 4 AAE07514 | AAe07514 Human PYR |
| 24 | 2396 | 43.8 | 1034 | 4 ABU08503 | ABU08503 Huma PYR |
| 25 | 2396 | 43.8 | 1034 | 6 ABU63315 | ABu63315 Human pyr |

ALIGNMENTS

RESULT 1
ADE36417

ID ADE36417 standard; protein; 1035 AA.

XX AC ADE36417;

XX XX 29-JAN-2004 (first entry)

XX DE Human PAA and nucleotide binding protein PAN6.

XX KW cytosolic; immunosuppressive; vulnary; antinflammatory; vasotropic;
KW antiallergic; antulcer; dermatological; cerebroprotective; cardiant;
KW antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
KW NF-kappaB activation inhibitor; PAA domain containing polypeptide;
KW PAA and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
KW apoptosis-associated spect-like protein; caspase recruitment domain 2;
KW ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
KW apoptosis; NF-kappaB induction; cytokine processing;
KW cytokine receptor signaling caspase-mediated proteolysis;
KW c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
KW inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
KW neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
KW fibrosis; smooth muscle cell proliferation; balloon angioplasty;
KW restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
KW arthritis; lupus; schrojen's syndrome; Crohn's disease;
KW ulcerative colitis; graft versus host disease; stroke; heart failure;
KW neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
KW cancer therapy; PAA domain family; human; PAN6.

XX OS Homo sapiens.

XX XX US2003077699-A1.

XX PD 24-APR-2003.

XX PF 25-SEP-2001; 2001US-00965621.

XX XX (GODZIK A.

PR (CHUZ) CHU Z.

PR PA (PAWL) PAWLOWSKI K.

PR PA (FIOR) FIORENTINO L.

PR PA (ARIZ) ARIZA M E.

XX PA (STEH) STEHLIK C.

XX XX

Wed Feb 9 10:22:14 2005

PI Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;
 PI Stehlik C;
 XX
 XX
 XX WPI: 2002-471256/50.
 DR N-PSDB; ADE36416.
 XX
 PT Novel isolated PAAD domain containing polypeptide useful for inducing
 PT apoptosis by inhibiting nuclear factor kappa B activation and in gene
 PT therapy for treating cancer.
 XX
 XX Claim 18; SEQ ID NO 24; 93pp; English.
 XX
 XX The invention describes an isolated PAAD domain containing polypeptide
 CC (1) comprising 80% identity to the amino acid sequence of PAAD and
 CC nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck
 CC -like protein containing a caspase recruitment domain (ASC)-2 fully
 CC defined in specification, where (1) is biologically active. (1) is useful
 CC for identifying a (1)-associated polypeptide, an agent altering that
 CC association and agents that modulate PAAD domain mediated inhibition of
 CC nuclear factor kappa B (NF-kappaB). A NB-ARC domain polypeptide is useful
 CC for identifying an agent that modulates the activity of the NB-ARC domain
 CC of (1). (1) or its functional fragments is useful in altering cellular or
 CC biochemical process such as apoptosis, NF-kappaB induction, cytokine
 CC processing, cytokine receptor signaling caspase-mediated proteolysis or c
 CC -Jun N-terminal kinase activation, thus having modulating effect on cell
 CC life and death (apoptosis) inflammation, cell adhesion or other cellular
 CC or biochemical processes. (1) is useful for treating cancer pathologies,
 CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
 CC hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
 CC proliferation in arteries following balloon angioplasty (restenosis),
 CC leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
 CC lupus, schrofen's syndrome, Crohn's disease and ulcerative colitis, graft
 CC versus host disease, stroke, heart failure, neurodegenerative diseases
 CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
 CC infection (HIV). (1) is useful for diagnosing cancer or monitoring cancer
 CC therapy. This is the amino acid sequence of a human PAAD and nucleotide
 CC binding protein PAN6.
 XX
 SQ Sequence 1035 AA;
 Query Match 100.0%; Score 5472; DB 5; Length 1035;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLRTAGDGLCLSTYLBELAEVAVELKFKLYLGATATLCEGKIPWGSMEKAGPLEMAQLL 60
 DB 1 MLRTAGDGLCLSTYLBELAEVAVELKFKLYLGATATLCEGKIPWGSMEKAGPLEMAQLL 60
 QY 61 ITHFGPEAWRLALSTFERINKDLWERGQREDLVRDPOETRYDYVRKFKRLMEDRNARL 120
 DB 61 ITHFGPEAWRLALSTFERINKDLWERGQREDLVRDPOETRYDYVRKFKRLMEDRNARL 120
 QY 121 GECVNLSHRYTRLLLVKHSNPMVQVQQLDTGRGHARTVGHQASPIKIETLFPDDEBP 180
 DB 121 GECVNLSHRYTRLLLVKHSNPMVQVQQLDTGRGHARTVGHQASPIKIETLFPDDEBP 180
 QY 181 EPPRTVMQGAAGIGKSMALAHKVMLDWADGKLFQGRFDYLFYINCREMNQATSCSMQDL 240
 DB 181 EPPRTVMQGAAGIGKSMALAHKVMLDWADGKLFQGRFDYLFYINCREMNQATSCSMQDL 240
 QY 241 IFSCWPEPSAPLQELIRVERLLFIIDGFDLKPSPHDPQGPWCICWEKKEPTELLNSL 300
 DB 241 IFSCWPEPSAPLQELIRVERLLFIIDGFDLKPSPHDPQGPWCICWEKKEPTELLNSL 300
 QY 301 IRKLLAPELSLITTRPTALEKHLLEHPRHVEILGFSEAEKKEYFYKPHNAOAGOV 360
 DB 301 IRKLLAPELSLITTRPTALEKHLLEHPRHVEILGFSEAEKKEYFYKPHNAOAGOV 360
 QY 361 FNVVRNEPLFTWCFCVPLVCVVCTCQQLGGGLLROTSRTTAVVYMLYLLSLMQPKP 420
 DB 361 FNVVRNEPLFTWCFCVPLVCVVCTCQQLGGGLLROTSRTTAVVYMLYLLSLMQPKP 420
 QY 421 GAPRLQPPNPQRLCSLAADGLWNQKILFEEQDLRKHGDLGSDVSAFLNMNIFQKDINCE 480

Db 421 GAPRLQPPNPQRLCSLAADGLWNQKILFEEQDLRKHGDLGSDVSAFLNMNIFQKDINCE 480
 QY 481 RYYSFIHLSFOEPFAAMYIILDEGEGGAGPDQDVTLLITYAFSERSFLALTSTRFLGLL 540
 Db 481 RYYSFIHLSFOEPFAAMYIILDEGEGGAGPDQDVTLLITYAFSERSFLALTSTRFLGLL 540
 QY 541 NEETRSHLEKSLCWKSPHIMDLQWIOSKAQSDGSTLQOQSGLEFFSCLYEIQEEEFIQ 600
 Db 541 NEETRSHLEKSLCWKSPHIMDLQWIOSKAQSDGSTLQOQSGLEFFSCLYEIQEEEFIQ 600
 QY 601 QALSHFQIVVSNIASKMEHVVSCFLKRCRSAQVLLHYGATYSADGEDRARCAGATHL 660
 Db 601 QALSHFQIVVSNIASKMEHVVSCFLKRCRSAQVLLHYGATYSADGEDRARCAGATHL 660
 QY 661 LVQLRPERTVLLDAYSEHLAAALCTNPNIETSLYRNALGSRGVKLLCOGLRHPNCKLQ 720
 Db 661 LVQLRPERTVLLDAYSEHLAAALCTNPNIETSLYRNALGSRGVKLLCOGLRHPNCKLQ 720
 QY 721 LRLKCRISSEACEDLSAALIANKNLTMDLSGNGVGPFGMMLLCEGLRHPQCRLOMIQL 780
 Db 721 LRLKCRISSEACEDLSAALIANKNLTMDLSGNGVGPFGMMLLCEGLRHPQCRLOMIQL 780
 QY 781 RKQLESAGCQEMASVLTGNPHLVELDITGNALDELGLRLCOGLRHPVCELRILWLKIC 840
 Db 781 RKQLESAGCQEMASVLTGNPHLVELDITGNALDELGLRLCOGLRHPVCELRILWLKIC 840
 QY 841 RLTAACDELASTLSVNSQSLRELDLSNELGDLVLLCEGLRHPCTCKLTIRLIGICRLG 900
 Db 841 RLTAACDELASTLSVNSQSLRELDLSNELGDLVLLCEGLRHPCTCKLTIRLIGICRLG 900
 QY 901 SAACGLSVVLOANHNRELDLSFNDLGDWGLWLLAEGLQHPACRLOKMLDSCGLTAKA 960
 Db 901 SAACGLSVVLOANHNRELDLSFNDLGDWGLWLLAEGLQHPACRLOKMLDSCGLTAKA 960
 QY 961 CENLYFTLGINTLTDLYLTNNALGDTGVRLCKELSHPGCKLRVWLFGMDLNKWTSHR 1020
 Db 961 CENLYFTLGINTLTDLYLTNNALGDTGVRLCKELSHPGCKLRVWLFGMDLNKWTSHR 1020
 QY 1021 LAALRVTKPYLDIGC 1035
 Db 1021 LAALRVTKPYLDIGC 1035
 RESULT 2
 ADJ19339
 ID ADJ19339 standard; protein; 1035 AA.
 XX
 AC ADJ19339;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Human PAN6 PAAD domain-containing full-length protein.
 XX
 KW PAAD; cytostatic; cancer; gene therapy; PAAD domain; human; PAN6.
 XX
 OS Homo sapiens.
 XX
 PN US2004002593-A1.
 XX
 PD 01-JAN-2004.
 XX
 PF 04-APR-2003; 2003US-00407866.
 XX
 PF 04-APR-2002; 2002US-0370538P.
 PR
 XX (REED/) REED J C.
 FA (GODZ/) GODZIK A.
 PI
 XX Reed JC, Godzik A;
 XX WPI: 2004-061677/06.
 DR N-PSDB; ADJ19338.
 DR

| | | | | | | | | | | | | | | | | | | | | |
|---|--|---|-----|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
| XX | New nucleic acid molecule encoding a PAAD-containing polypeptide, useful for preparing a composition for diagnosing or treating e.g., cancer. | | | | | | | | | | | | | | | | | | | |
| PT | Example; SEQ ID NO 24; 193pp; English. | | | | | | | | | | | | | | | | | | | |
| PS | The invention relates to a novel isolated nucleic acid molecule encoding a PAAD-containing polypeptide. The polypeptide of the invention demonstrates cytostatic activity and may be useful for preparing a composition for diagnosing or treating diseases associated with the PAAD domain-containing polypeptide, such as cancer, via gene therapy. The current sequence is that of the human PAAD domain-containing full-length protein of the invention. | | | | | | | | | | | | | | | | | | | |
| XX | SQ Sequence 1035 AA; | | | | | | | | | | | | | | | | | | | |
| Query Match 100.0%; Score 5472; DB 8; Length 1035; | | | | | | | | | | | | | | | | | | | | |
| Best Local Similarity 100.0%; Pred. No. 0; | | | | | | | | | | | | | | | | | | | | |
| Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | | | | | | | | | | | | |
| QY | 1 | MLRTAGDGLCRSLTYLBEAVALKFKLYLGTATLGEKGKIPWGSMEKAGPLEMAQLL | 60 | | | | | | | | | | | | | | | | | |
| DB | 1 | MLRTAGDGLCRSLTYLBEAVALKFKLYLGTATLGEKGKIPWGSMEKAGPLEMAQLL | 60 | | | | | | | | | | | | | | | | | |
| QY | 61 | ITHFGPEAWRLALSTFERINRKLWERGQREDLVDPQETVRYDYRRKFRIMEDRNARL | 120 | | | | | | | | | | | | | | | | | |
| DB | 61 | ITHFGPEAWRLALSTFERINRKLWERGQREDLVDPQETVRYDYRRKFRIMEDRNARL | 120 | | | | | | | | | | | | | | | | | |
| QY | 121 | GEVNLSHRYTRLLLVKSHSNPMQVQQQLDTGRGHARTVGHQASPIKIETLFPDEBRP | 180 | | | | | | | | | | | | | | | | | |
| DB | 121 | GEVNLSHRYTRLLLVKSHSNPMQVQQQLDTGRGHARTVGHQASPIKIETLFPDEBRP | 180 | | | | | | | | | | | | | | | | | |
| QY | 181 | EPPTVVMQGAAGICKSLAHKVMWLDWADGKLFQGRFDYLYFVINCENMQSATECSMDL | 240 | | | | | | | | | | | | | | | | | |
| DB | 181 | EPPTVVMQGAAGICKSLAHKVMWLDWADGKLFQGRFDYLYFVINCENMQSATECSMDL | 240 | | | | | | | | | | | | | | | | | |
| QY | 241 | IFSCWPEPSAPLQELIRVPERLLFTIDGFDELKPSFHDPPQGPWCLCWEKRPTELLNSL | 300 | | | | | | | | | | | | | | | | | |
| DB | 241 | IFSCWPEPSAPLQELIRVPERLLFTIDGFDELKPSFHDPPQGPWCLCWEKRPTELLNSL | 300 | | | | | | | | | | | | | | | | | |
| QY | 301 | IRKLLLPESLLITTRPTALEKHLRHPHVEILGFSEAEKKEYFYKYFHNAEQAGQV | 360 | | | | | | | | | | | | | | | | | |
| DB | 301 | IRKLLLPESLLITTRPTALEKHLRHPHVEILGFSEAEKKEYFYKYFHNAEQAGQV | 360 | | | | | | | | | | | | | | | | | |
| QY | 361 | FNYYRDNPEFTMCFVPLVWVCTCLOQQLGEGGLLRQTSRTTTAVYMLYLLSLMQPKP | 420 | | | | | | | | | | | | | | | | | |
| DB | 361 | FNYYRDNPEFTMCFVPLVWVCTCLOQQLGEGGLLRQTSRTTTAVYMLYLLSLMQPKP | 420 | | | | | | | | | | | | | | | | | |
| QY | 421 | GAPRLQPPNQRGLCSLAADGLWNQKILFEEQDLRKHGLDGEDVSAFLNMNIFQKDINCE | 480 | | | | | | | | | | | | | | | | | |
| DB | 421 | GAPRLQPPNQRGLCSLAADGLWNQKILFEEQDLRKHGLDGEDVSAFLNMNIFQKDINCE | 480 | | | | | | | | | | | | | | | | | |
| QY | 481 | RYYSFTHLSFOEFPFAAMYIILDEGGGAGPDQDVTRLLTEYAFSERSFLATSTFLGILL | 540 | | | | | | | | | | | | | | | | | |
| DB | 481 | RYYSFTHLSFOEFPFAAMYIILDEGGGAGPDQDVTRLLTEYAFSERSFLATSTFLGILL | 540 | | | | | | | | | | | | | | | | | |
| QY | 541 | NEETRSHLEKSLCWKVSPIHIMDLQWITQSKAQSDGSTLQOQSLEFFSCLYEIEDEEFIQ | 600 | | | | | | | | | | | | | | | | | |
| DB | 541 | NEETRSHLEKSLCWKVSPIHIMDLQWITQSKAQSDGSTLQOQSLEFFSCLYEIEDEEFIQ | 600 | | | | | | | | | | | | | | | | | |
| QY | 601 | QALSHFQVIVVNSITASKMEHVMVSSFLCKRCRQAQVILHYGATYSADGEDRARCAGATL | 660 | | | | | | | | | | | | | | | | | |
| DB | 601 | QALSHFQVIVVNSITASKMEHVMVSSFLCKRCRQAQVILHYGATYSADGEDRARCAGATL | 660 | | | | | | | | | | | | | | | | | |
| QY | 661 | LVQLRPERTVLLDAYSEHLAAALCTNPNLIELSLYRNALSGRGVKLLCOGLRHHPNCKLQN | 720 | | | | | | | | | | | | | | | | | |
| DB | 661 | LVQLRPERTVLLDAYSEHLAAALCTNPNLIELSLYRNALSGRGVKLLCOGLRHHPNCKLQN | 720 | | | | | | | | | | | | | | | | | |
| QY | 721 | LRLKRCRISSSACEDLSAALIANKNLTMRDLSGNVGVPFGWMLLCEGLRHHPQCRLOMTQL | 780 | | | | | | | | | | | | | | | | | |
| DB | 721 | LRLKRCRISSSACEDLSAALIANKNLTMRDLSGNVGVPFGWMLLCEGLRHHPQCRLOMTQL | 780 | | | | | | | | | | | | | | | | | |
| QY | 781 | RKQLESAGACQEMASVLGTNPVHVELDLTGNALEDGLRLLCQGLRHPVCRRLTLWLKIC | 840 | | | | | | | | | | | | | | | | | |

| | | | |
|----------|---|--|------|
| Db | 781 | RKQLESAGACQEMASVLGTNPVHVELDTGNALDGLRLLCQGLRHPVCRRLTLWLKIC | 840 |
| Qy | 841 | RLTAAACDELAASLVNQSLRELDLSNELGDLGVLLLCGLRHPTCKLQTLRLGICRLG | 900 |
| Db | 841 | RLTAAACDELAASLVNQSLRELDLSNELGDLGVLLLCGLRHPTCKLQTLRLGICRLG | 900 |
| Qy | 901 | SAACEGLSVVLQANHNLRDLSPNDLGDWGLWLLAEGLQHPACRLQKWLWDSGLTAKA | 960 |
| Db | 901 | SAACEGLSVVLQANHNLRDLSPNDLGDWGLWLLAEGLQHPACRLQKWLWDSGLTAKA | 960 |
| Qy | 961 | CENLYFTLGINQTLTDLVLTNNALGDTGVRLLCRLSHPGCKRLVWLFQMDLNKMTSHR | 1020 |
| Db | 961 | CENLYFTLGINQTLTDLVLTNNALGDTGVRLLCRLSHPGCKRLVWLFQMDLNKMTSHR | 1020 |
| Qy | 1021 | LAALRVTKPYLDIGC | 1035 |
| Db | 1021 | LAALRVTKPYLDIGC | 1035 |
| RESULT 3 | | | |
| ADP47724 | | | |
| ID | ADP47724 standard; protein; 1062 AA. | | |
| XX | ADP47724; | | |
| XX | 12-AUG-2004 (first entry) | | |
| XX | Human Monarch-1 (CATERPILLER 19.3) full-length protein. | | |
| DE | Monarch-1; CATERPILLER 11.2; caspase recruitment domain; | | |
| KW | CARD, transcription enhancer, R-binding, pyrin, lots of leucine repeat; | | |
| KW | CATERPILLER 11.3; CATERPILLER 16.1; CATERPILLER 16.2; CIAS1; | | |
| KW | cold-induced autoinflammatory syndrome 1; antinflammatory; cytostatic; | | |
| KW | inflammatory disease; cancer; gene therapy; human; purine; | | |
| KW | CATERPILLER 19.3; chromosome 19q13; multiple sclerosis. | | |
| OS | Homo sapiens. | | |
| XX | WO2004034093-A2. | | |
| XX | 22-APR-2004. | | |
| XX | 30-APR-2003; 2003WO-US013562. | | |
| XX | 30-APR-2002; 2002US-0376626P. | | |
| XX | (UYNC-) UNIV NORTH CAROLINA. | | |
| XX | Ting JY, Linhoff MW, Harton JA, Williams KL, Lich J, O'Connor W; | | |
| PI | Moore CB, Davis B, Brickey J, Conti B, Zhang J, Zhu X; | | |
| XX | WPI; 2004-348215/32. | | |
| DR | N-FSDB; ADP47723. | | |
| PT | New nucleic acid encoding Monarch-1, CATERPILLER 11.2, CATERPILLER 11.3, | | |
| PT | CATERPILLER 16.1, CATERPILLER 16.2 or CIAS1 polypeptide, useful in | | |
| XX | preparing a composition for treating inflammatory disease or cancer. | | |
| XX | Claim 18; SEQ ID NO 2; 205pp; English. | | |
| XX | The invention relates to a novel isolated nucleic acid encoding a Monarch | | |
| CC | -1, CATERPILLER (CARD [caspase recruitment domain], transcription | | |
| CC | enhancer, R(purine)-binding, pyrin, lots of leucine repeats) 11.2, | | |
| CC | CATERPILLER 11.3, CATERPILLER 16.1, CATERPILLER 16.2 or CIAS1 (cold- | | |
| CC | induced autoinflammatory syndrome 1) polypeptide comprising the amino | | |
| CC | acids encoded by exon 6 and lacking the amino acids encoded by exon 4 or | | |
| CC | its fragment. The nucleic acid of the invention demonstrates | | |
| CC | antinflammatory and cytostatic activities and may be useful in preparing | | |
| CC | a composition for treating an inflammatory disease or cancer, possibly | | |
| CC | via gene therapy. The current sequence is that of the human Monarch-1 | | |
| CC | (CATERPILLER 19.3) full-length protein of the invention which is encoded | | |
| CC | by DNA located on chromosome 19q13, in the multiple sclerosis | | |

Qy 1 MLRTAGDGLCRSLTYLSELEAVELKKFKLYLGTATLGECKIPWGSMEKAGPLEMAQLL 60
Db 1 MLRTAGDGLCRSLTYLSELEAVELKKFKLYLGTATLGECKIPWGSMEKAGPLEMAQLL 60
Qy 61 ITHFGPEEAWRLALSTFERINRKLWEGQEDLVR----- 96
Db 61 ITHFGPEEAWRLALSTFERINRKLWEGQEDLVR----- 120
Qy 97 ---DPQETRYDYRRKFRMLMDNRALGECVNLSHRYTRLLLVKEHSPMPVQOQLDGTG 153
Db 121 PKQPOETRYDYRRKFRMLMDNRALGECVNLSHRYTRLLLVKEHSPMPVQOQLDGTG 180
Qy 154 RGHARTVGHQASPIKIEFLFPDEERPPPTVVMQGAAGIKGKMLAHKVMLDWADGKLF 213
Db 181 RGHARTVGHQASPIKIEFLFPDEERPPPTVVMQGAAGIKGKMLAHKVMLDWADGKLF 240
Qy 214 QGRDYLYINCREMNOGATSCQDLIFSCWPEPSAPLOELIRVPERLLFIIDGFDL 273
Db 241 QGRDYLYINCREMNOGATSCQDLIFSCWPEPSAPLOELIRVPERLLFIIDGFDL 300
Qy 274 PSFHDPOQWMLCWEKRPTELLNSLRKLLPELSLITRPTALEKHLRLEHPRHV 333
Db 301 PSFHDPOQWMLCWEKRPTELLNSLRKLLPELSLITRPTALEKHLRLEHPRHV 360
Qy 334 EILGFSEAEKKEYFYKFNHAEQAGQVFNVDNEPLFTMCFVPLVCWVCTCLQQOLEG 393
Db 361 EILGFSEAEKKEYFYKFNHAEQAGQVFNVDNEPLFTMCFVPLVCWVCTCLQQOLEG 420
Qy 394 GGLLRQTSRTTAVMYLLSLMOPKQAPLQPPNQRLGCLSAADGLMNQKILFREQD 453
Db 421 GGLLRQTSRTTAVMYLLSLMOPKQAPLQPPNQRLGCLSAADGLMNQKILFREQD 480
Qy 454 LRKGLDGEDVSALFNWNI FQDINCERYYSFIHLSPQFAAMYIILDEGEGAGDQD 513
Db 481 LRKGLDGEDVSALFNWNI FQDINCERYYSFIHLSPQFAAMYIILDEGEGAGDQD 540
Qy 514 VTRLLTEYAFSERSFLATSRFLGLLNEETRSLEKSLCKVSPHIMDLLOWIQSKAQ 573
Db 541 VTRLLTEYAFSERSFLATSRFLGLLNEETRSLEKSLCKVSPHIMDLLOWIQSKAQ 600
Qy 574 SDGSTLOQSLSEFFSCLYEIEBEFFIQALSHFQVIVVNSIASQVHEWVSFCLKRCR 633
Db 601 SDGSTLOQSLSEFFSCLYEIEBEFFIQALSHFQVIVVNSIASQVHEWVSFCLKRCR 660
Qy 634 QVLHLYGATYSADGEDRARCAGATLLVQLRPTVLLDAYSEHLAAALCTNPNLIELS 693
Db 661 QVLHLYGATYSADGEDRARCAGATLLVQLRPTVLLDAYSEHLAAALCTNPNLIELS 719
Qy 694 LYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSG 753
Db 720 LYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSG 779
Qy 754 NGVGFPGMMLICEGLRHPQRLQMLQRLKCOLESAGQEMASVLTGNPHLVLDLTGNAL 813
Db 780 NGVGFPGMMLICEGLRHPQRLQMLQRLKCOLESAGQEMASVLTGNPHLVLDLTGNAL 839
Qy 814 EDLGLRLCCQGLRHPVCLRLTLWKICRLTAAACDELASTLSVNSQSLRELDLSNELGDL 873
Db 840 EDLGLRLCCQGLRHPVCLRLTLWKICRLTAAACDELASTLSVNSQSLRELDLSNELGDL 899
Qy 874 GVLLICEGLRHPTCKLQTLRLGICRLGSAACEGLSVVLQANHLRELDLSFNDLGDWGLW 933
Db 900 GVLLICEGLRHPTCKLQTLRLGICRLGSAACEGLSVVLQANHLRELDLSFNDLGDWGLW 959
Qy 934 LLAEGLOHPACRLQKMLDSCGLTAKACENIYFTLGINQTLTLDYLTNNALGDTGVRLLC 993
Db 960 LLAEGLOHPACRLQKMLDSCGLTAKACENIYFTLGINQTLTLDYLTNNALGDTGVRLLC 1019
Qy 994 KRLSHPGCKLVLVFGMDLNKMTHTSLAALRVTKPYLDIGC 1035
Db 1020 KRLSHPGCKLVLVFGMDLNKMTHTSLAALRVTKPYLDIGC 1061

RESULT 5
ADF94755
ID ADF94755 standard; protein; 1061 AA.
XX
AC ADF94755;
XX
DT 26-FEB-2004 (first entry)
XX
DB Human PYRIN-8 protein #2.
XX
KW human; PYRIN; inflammatory disorder; inappropriate apoptosis;
KW inflammatory bowel disease; rheumatoid arthritis; diabetes;
KW multiple sclerosis; Grave's disease; contact dermatitis; psoriasis;
KW graft rejection; asthma; allergy; chronic obstructive pulmonary disease;
KW glomerulonephritis; infection; Alzheimer's disease; Parkinson's disease;
KW anaemia; ischaemia; screening; chromosomal mapping; tissue typing;
KW forensic biology; pharmacogenomics; predictive medicine.
XX
OS Homo sapiens.
XX
PN WO2003089588-A2.
XX
PD 30-OCT-2003.
XX
PF 14-APR-2003; 2003WO-US011572.
XX
PR 17-APR-2002; 2002US-00124498.
XX
PA (MILL-) MILLENNIUM PHARM INC.
PA (AMHP) WYETH.
XX
PI Bertin J, Wang W, Blatcher M;
XX
DR WPI; 2003-845527/78.
DR N-PSDB; ADF94754.
XX
XX
XX New nucleic acid molecules and polypeptides (e.g. PYRIN-2 or PYRIN-3)
XX useful for diagnosing, preventing or treating inflammation or disorders
XX associated with inappropriate apoptosis, in chromosomal mapping or in
XX pharmacogenomics.
XX
PS Claim 8; SEQ ID NO 18; 199pp; English.
XX
CC The invention comprises the amino acid and coding sequences of human
CC PYRIN proteins. The DNA and protein sequences of the invention are useful
CC in diagnosing, preventing and treating inflammatory disorders or
CC disorders associated with inappropriate apoptosis, such as: inflammatory
CC bowel disease, rheumatoid arthritis, diabetes, multiple sclerosis,
CC Grave's disease, contact dermatitis, psoriasis, graft rejection, asthma,
CC allergy, chronic obstructive pulmonary disease, glomerulonephritis,
CC infections, Alzheimer's disease, Parkinson's disease, anaemia and
CC ischaemia. The DNA and protein sequences of the invention may also be
CC used in screening assays, chromosomal mapping, tissue typing, forensic
CC biology, pharmacogenomics, predictive medicine, and in monitoring of
CC clinical trials. The present amino acid sequence represents a PYRIN
CC protein of the invention.
XX
SQ Sequence 1061 AA;

Query Match 99.3%; Score 5433; DB 7; Length 1061;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1034; Conservative 0; Mismatches 0; Indels 28; Gaps 2;
Qy 1 MLRTAGDGLCRSLTYLSELEAVELKKFKLYLGTATLGECKIPWGSMEKAGPLEMAQLL 60
Db 1 MLRTAGDGLCRSLTYLSELEAVELKKFKLYLGTATLGECKIPWGSMEKAGPLEMAQLL 60
Qy 61 ITHFGPEEAWRLALSTFERINRKLWEGQEDLVR----- 96
Db 61 ITHFGPEEAWRLALSTFERINRKLWEGQEDLVR----- 120
Qy 97 ---DPQETRYDYRRKFRMLMDNRALGECVNLSHRYTRLLLVKEHSPMPVQOQLDGTG 153

QY 361 FNYVRDNEPLFTMCFVPLVWVCTCQQQLEGGGLLRQTSRTTTAVYMLYLLSLMQPKP 420
Db |||||
361 FNYVRDNEPLFTMCFVPLVWVCTCQQQLEGGGLLRQTSRTTTAVYMLYLLSLMQPKP 420
QY 421 GAPRLQPPNORGICSLAAGLWQKILFEEQDLRKXGLDGEDVSAFLNNNIFOKDINCE 480
Db |||||
421 GAPRLQPPNORGICSLAAGLWQKILFEEQDLRKXGLDGEDVSAFLNNNIFOKDINCE 480
QY 481 RYYSFIHLSFOEFFAAMYIILDEEGGAGPQDVTRLLITEYAFSERSFLALTSRFLGCLL 540
Db |||||
481 RYYSFIHLSFOEFFAAMYIILDEEGGAGPQDVTRLLITEYAFSERSFLALTSRFLGCLL 540
QY 541 NEETRSHLEKSLCWKVSPIHKDILLQWTSQAQSDGSTLQOGSLEFFSCLYEIQEEFFIQ 600
Db |||||
541 NEETRSHLEKSLCWKVSPIHKDILLQWTSQAQSDGSTLQOGSLEFFSCLYEIQEEFFIQ 600
QY 601 QALSHFQIVVNSNTASKMEHVVSSFCCKRCSAQLVHLHGATYSADGEDRARCAGATL 660
Db |||||
601 QALSHFQIVVNSNTASKMEHVVSSFCCKRCSAQLVHLHGATYSADGEDRARCAGATL 660
QY 661 LVQLRPERTVLLDAYSEHLAALCTNPNLIELSLYRNALSGRGVKKLLCOGLRHPNCKLQN 720
Db |||||
661 LVQLRPERTVLLDAYSEHLAALCTNPNLIELSLYRNALSGRGVKKLLCOGLRHPNCKLQN 720
QY 721 LRLKCRISACBDLSAALIANKNLTRMDLSGNGVGPFGMWLLCEGLRHPQCRLOMIQL 780
Db |||||
721 LRLKCRISACBDLSAALIANKNLTRMDLSGNGVGPFGMWLLCEGLRHPQCRLOMIQL 780
QY 781 RKCOLESGACQEMASVLGTNPHVELDITGNALDGLRLICQGLRHPVCELRITLWLKIC 840
Db |||||
781 RKCOLESGACQEMASVLGTNPHVELDITGNALDGLRLICQGLRHPVCELRITLWLKIC 840
QY 841 RLTAACDELASTLSVNOISRELDSLNELDGLVLLICEGLRHPCTCKLQTLRLGICRLG 900
Db |||||
841 RLTAACDELASTLSVNOISRELDSLNELDGLVLLICEGLRHPCTCKLQTLRLGICRLG 900
QY 901 SAACEGLSVLQANHNLRELDSLNFNDLGDWGLWLLAEGLOHPACRLQKLMWDSGLTAKA 960
Db |||||
901 SAACEGLSVLQANHNLRELDSLNFNDLGDWGLWLLAEGLOHPACRLQKLMWDSGLTAKA 960
QY 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLICKRLSHPGCKLRVLMFGMDLNKMTS 1019
Db |||||
961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLICKRLSHPGCKLRVLMFGMDLNKMTS 1019

RESULT 7

ADJ19383
ID ADJ19383 standard; protein; 1027 AA.
XX
AC ADJ19383;
XX
AC ADJ19383;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human PAN6 PAAD domain-containing full-length protein 2.
XX
KW PAAD; cytostatic; cancer; gene therapy; human; PAAD domain; PAN6.
XX
OS Homo sapiens.
XX
FN US2004002593-A1.
XX
PD 01-JAN-2004.
XX
PF 04-APR-2003; 2003US-00407866.
XX
PR 04-APR-2002; 2002US-0370538P.
XX
PA (REED/) REED J C.
XX
PA (GODZ/) GODZIK A.
XX
PI Reed JC, Godzik A;
XX
DR WPI; 2004-061677/06.

DR N-PSDB; ADJ19382.
XX
PT New nucleic acid molecule encoding a PAAD-containing polypeptide, useful
PT for preparing a composition for diagnosing or treating e.g., cancer.
XX
PS Claim 14; SEQ ID NO 68; 193pp; English.
XX
CC The invention relates to a novel isolated nucleic acid molecule encoding
CC a PAAD-containing polypeptide. The polypeptide of the invention
CC demonstrates cytostatic activity and may be useful for preparing a
CC composition for diagnosing or treating diseases associated with the PAAD
CC domain-containing polypeptide, such as cancer, via gene therapy. The
CC current sequence is that of the human PAAD domain-containing full-length
CC protein of the invention.
XX
SQ Sequence 1027 AA;

Query Match 96.1%; Score 5259.5; DB 8; Length 1027;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1000; Conservative 0; Mismatches 0; Indels 27; Gaps 1;
QY 1 MLRTAGRDGLCRSLSTYLEELEAVELEKPKLYLGTATLGEKGKIPWGSMEKAGPLEMAQLL 60
Db |||||
1 MLRTAGRDGLCRSLSTYLEELEAVELEKPKLYLGTATLGEKGKIPWGSMEKAGPLEMAQLL 60
QY 61 ITHFGPEANRLALSTFERINRKDLWERGQREDLVR----- 96
Db |||||
61 ITHFGPEANRLALSTFERINRKDLWERGQREDLVRTPPGPSSSLGNQSTCLLEVSILVT 120
QY 97 ---DPQSTYDYVRRKPRLMEDNRNARLGEVCNLSHRYTRLLLVKEHNPMPVQQQLDGT 153
Db |||||
121 PRKDPQSTYDYVRRKPRLMEDNRNARLGEVCNLSHRYTRLLLVKEHNPMPVQQQLDGT 180
QY 154 RGHARTVGHQASPIKTIETLPEDEERPEPRTVVMQGAAGIGKSMIAHKWMLDWDGKLF 213
Db |||||
181 RGHARTVGHQASPIKTIETLPEDEERPEPRTVVMQGAAGIGKSMIAHKWMLDWDGKLF 240
QY 214 QGRFDYLYFYNCREMNSQATECSMDLI FSCWPEPSAPLOELIRVPERLLFIIDGFDELK 273
Db |||||
241 QGRFDYLYFYNCREMNSQATECSMDLI FSCWPEPSAPLOELIRVPERLLFIIDGFDELK 300
QY 274 PSFHDPOGWCLEWEKRPTELLNLSIRKKLPDELSTLTRPTALEKHLRLEHPRHV 333
Db |||||
301 PSFHDPOGWCLEWEKRPTELLNLSIRKKLPDELSTLTRPTALEKHLRLEHPRHV 360
QY 334 EILGFSEAEKKEYFYKFNHNAEQAGVFNVRNEPLFTMCFVPLVWVCTCQQQLEGG 393
Db |||||
361 EILGFSEAEKKEYFYKFNHNAEQAGVFNVRNEPLFTMCFVPLVWVCTCQQQLEGG 420
QY 394 GGLLRQTSRTTTAVYMLYLLSLMQPKPGAPRLQPPNORGICSLAAGLWQKILFEEQD 453
Db |||||
421 GGLLRQTSRTTTAVYMLYLLSLMQPKPGAPRLQPPNORGICSLAAGLWQKILFEEQD 480
QY 454 LRKHGLDGEDVSAFLNNNIFOKDINCERYYSFIHLSFOEFFAAMYIILDEEGGAGPQD 513
Db |||||
481 LRKHGLDGEDVSAFLNNNIFOKDINCERYYSFIHLSFOEFFAAMYIILDEEGGAGPQD 540
QY 514 VTRLLITEYAFSERSFLALTSRFLGGLNEETRSHLEKSLCWKVSPIHKDILLQWISKAQ 573
Db |||||
541 VTRLLITEYAFSERSFLALTSRFLGGLNEETRSHLEKSLCWKVSPIHKDILLQWISKAQ 600
QY 574 SDGSTLQOGSLEFFSCLYEIQEEFFIQOALSHFOIVVNSNTASKMEHVVSSFCCKRCSA 633
Db |||||
601 SDGSTLQOGSLEFFSCLYEIQEEFFIQOALSHFOIVVNSNTASKMEHVVSSFCCKRCSA 660
QY 634 QVLHLYGATYSADGEDRARCAGATLVLVLRPERTVLLDAYSEHLAALCTNPNLIELS 693
Db |||||
661 QVLHLYGATYSADGEDRARCAGATLVLVLRPERTVLLDAYSEHLAALCTNPNLIELS 720
QY 694 LYRNALSGRGVKKLLCOGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKLTRMDLSG 753
Db |||||
721 LYRNALSGRGVKKLLCOGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKLTRMDLSG 780

Qy 754 NGVFFPGMMLLCEGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 813
 Db 781 NGVFFPGMMLLCEGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 840
 Qy 814 EDGLRLCCGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 873
 Db 841 EDGLRLCCGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 900
 Qy 874 GVLLCEGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 933
 Db 901 GVLLCEGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 960
 Qy 934 LLAEGLOHPACRLQKWLDSOGLTAKACENLYFTLGINQTLTDLTLTNALGDTGVRLLC 993
 Db 961 LLAEGLOHPACRLQKWLDSOGLTAKACENLYFTLGINQTLTDLTLTNALGDTGVRLLC 1020
 Qy 994 KRLSHPG 1000
 Db 1021 KRLSHPG 1027

RESULT 8
 ADP47726
 ID ADP47726 standard; protein; 1006 AA.
 XX AC ADP47726;
 XX AC ADP47726;
 DT 12-AUG-2004 (first entry)
 XX Human Monarch-1 (CATERPILLER 19.3) isoform II protein.
 XX Monarch-1; CATERPILLER 11.2; caspase recruitment domain;
 KW CARD, transcription enhancer, R-binding, pyrin, lots of leucine repeat;
 KW CATERPILLER 11.3; CATERPILLER 16.1; CATERPILLER 16.2; CIASI; cold-
 KW cold-induced autoinflammatory syndrome 1; antiinflammatory; cytostatic;
 KW inflammatory disease; cancer; gene therapy; human; purine;
 KW CATERPILLER 19.3; chromosome 19q13; multiple sclerosis; isoform II.
 XX Homo sapiens.
 OS
 XX W02004034093-A2.
 PN 22-APR-2004.
 PD 30-APR-2003; 2003WO-US013562.
 PF 30-APR-2002; 2002US-0376626P.
 XX (UYNC-) UNIV NORTH CAROLINA.
 XX Ting JY, Linhoff MW, Harton JA, Williams KL, Lich J, O'connor W;
 PI Moore CB, Davis B, Brickey J, Conti B, Zhang J, Zhu X;
 XX WPI; 2004-348215/32.
 DR N-PSDB; ADP47725.
 XX New nucleic acid encoding Monarch-1, CATERPILLER 11.2, CATERPILLER 11.3,
 PT CATERPILLER 16.1, CATERPILLER 16.2 or CIASI polypeptide, useful in
 PT preparing a composition for treating inflammatory disease or cancer.
 XX Claim 18; SEQ ID NO 4; 205pp; English.
 PS The invention relates to a novel isolated nucleic acid encoding a Monarch
 CC -1, CATERPILLER (CARD [caspase recruitment domain], transcription
 CC enhancer, R(purine)-binding, pyrin, lots of leucine repeats) 11.2,
 CC CATERPILLER 11.3, CATERPILLER 16.1, CATERPILLER 16.2 or CIASI (cold-
 CC induced autoinflammatory syndrome 1) polypeptide comprising the amino
 CC acids encoded by exon 6 and lacking the amino acids encoded by exon 4 or
 CC its fragment. The nucleic acid of the invention demonstrates
 CC antiinflammatory and cytostatic activities and may be useful in preparing
 CC a composition for treating an inflammatory disease or cancer, possibly
 CC via gene therapy. The current sequence is that of the human Monarch-1
 CC (CATERPILLER 19.3) isoform II protein of the invention which is encoded

CC by DNA located on chromosome 19q13, in the multiple sclerosis
 XX susceptibility region.
 SQ Sequence 1006 AA;
 Query Match 93.4%; Score 5113.5; DB 8; Length 1006;
 Best Local Similarity 92.2%; Pred. No. 0;
 Matches 979; Conservative 0; Mismatches 0; Indels 83; Gaps 2;
 Qy 1 MLRTAGRDGLRLSTVLELEAVELKKFKLYLTATTELCGEGKIPWGSMEKAGLEMAQLL 60
 Db 1 MLRTAGRDGLRLSTVLELEAVELKKFKLYLTATTELCGEGKIPWGSMEKAGLEMAQLL 60
 Qy 61 ITHFGPEEARLALSTFERINRKLWERGOREDLVR----- 96
 Db 61 ITHFGPEEARLALSTFERINRKLWERGOREDLVRTPPGPSSIGNOSTCLLEVSLVT 120
 Qy 97 ----DPOETYRDYVRRKFRMLMEDRNARLGEVNLVSHRYTRLLLVKHSNPMQVQQLLDTG 153
 Db 121 PRKDPQETVYRDYVRRKFRMLMEDRNARLGEVNLVSHRYTRLLLVKHSNPMQVQQLLDTG 180
 Qy 154 RGHARTVGHQASPIKTIETLFEPEDEERPEPRTVMQGAAGIGKSMIAHKVMDLWDGKLF 213
 Db 181 RGHARTVGHQASPIKTIETLFEPEDEERPEPRTVMQGAAGIGKSMIAHKVMDLWDGKLF 240
 Qy 214 QGRFDYLFVINCENMQSATECSMODLIFSCWPEPSAPLOELIRVPERLLFIIDGFDELK 273
 Db 241 QGRFDYLFVINCENMQSATECSMODLIFSCWPEPSAPLOELIRVPERLLFIIDGFDELK 300
 Qy 274 PSFHPDQGPWCCLWEKRPTELLNSLIRKLLPELSLITTRPTALEKHLRLEHPRHV 333
 Db 301 PSFHPDQGPWCCLWEKRPTELLNSLIRKLLPELSLITTRPTALEKHLRLEHPRHV 360
 Qy 334 EILGFSEARKEYFYKYFNABQAGOVFNVDNEPLFTWCVPVLCVWVCTCLOOLEG 393
 Db 361 EILGFSEARKEYFYKYFNABQAGOVFNVDNEPLFTWCVPVLCVWVCTCLOOLEG 420
 Qy 394 GGLLRQTSRTTAVVYMLYLLSLMOPKPGAPRLOPPNQRGLCSLAADGLWNQKILFEEQD 453
 Db 421 GGLLRQTSRTTAVVYMLYLLSLMOPKPGAPRLOPPNQRGLCSLAADGLWNQKILFEEQD 480
 Qy 454 LRKHGLDGEDVSAFLNMIQKDIINERYYSFIHLSFQEFFFAAMYIILDEGEGAGPDQD 513
 Db 481 LRKHGLDGEDVSAFLNMIQKDIINERYYSFIHLSFQEFFFAAMYIILDEGEGAGPDQD 540
 Qy 514 VTRLLTEYAFSERSFLALTSRFLGLNEETSRLKSLCWKVSPIH KMDLLQMIQSKAQ 573
 Db 541 VTRLLTEYAFSERSFLALTSRFLGLNEETSRLKSLCWKVSPIH KMDLLQMIQSKAQ 600
 Qy 574 SDGSTLQOQSLEFFSCLYEIQEEFTQQALSHFQVTVVSNIAKSMHVMVSSFCLEKCRSA 633
 Db 601 SDGSTLQOQSLEFFSCLYEIQEEFTQQALSHFQVTVVSNIAKSMHVMVSSFCLEKCRSA 660
 Qy 634 QVLHLYGATYSADGEDRARCAGAHNTLVQLRPERVTLVDAYSEHAAALCTNPNIELS 693
 Db 661 QVLHLYGATYSADGEDRARCAGAHNTLVQLRPERVTLVDAYSEHAAALCTNPNIELS 720
 Qy 694 LYRNALGSRGVKLLCOGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 753
 Db 721 LYRNALGSRGVKLLCOGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 780
 Qy 754 NGVFFPGMMLLCEGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 813
 Db 781 NGVFFPGMMLLCEGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 840
 Qy 814 EDGLRLCCGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 873
 Db 841 EDGLRLCCGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 900
 Qy 874 GVLLCEGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 933
 Db 901 GVLLCEGLRHPCQLQMIQRKQLESAGCQEMASVLGTNPVHVELDLTGAL 960

QY 934 LLAEGLOHPACRLQKLVLDSCGLTAKACENLYFTLGINQTLTDLTLNNALGDTGVRLLC 993
 DB 961 LLAEGLOHPACRLQKLVLDSCGLTAKACENLYFTLGINQTLTDLTLNNALGDTGVRLLC 997
 QY 994 KRLSHPGCKLRVLMFGMDLNKMTSHSLAALRVTKPYLDIGC 1035
 DB 978 -----WLFMGMDLNKMTSHSLAALRVTKPYLDIGC 1006

RESULT 9

ADM04476
 ID ADM04476 standard; protein; 1004 AA.
 AC ADM04476;
 DT 20-MAY-2004 (first entry)
 DE Human protein of the invention SEQ ID NO:3161.
 KW human; gene therapy; diagnostic marker; pharmaceutical.
 OS Homo sapiens.
 PN EPI347046-A1.
 XX 24-SEP-2003.
 PF 12-APR-2002; 2002EP-00008400.
 PR 22-MAR-2002; 2002JP-00137785.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Negahari K, Masuho Y;
 DR WPI; 2003-723558/69.
 DR N-PSDB; ADM02033.
 XX New polynucleotides and polypeptides are useful in gene therapy, for
 PT developing a diagnostic marker or medicines for regulating their
 PT expression and activity, or as a target of gene therapy.
 XX Claim 1; SEQ ID NO 3161; 305pp; English.
 CC The invention relates to a novel human polynucleotide and the encoded
 CC polypeptide. A polynucleotide of the invention may have a use in gene
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
 CC as a primer for synthesizing the polynucleotide or as a probe for
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
 CC useful in gene therapy, for developing a diagnostic marker or medicines
 CC for regulating their expression and activity, or as a target of gene
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
 CC are useful as pharmaceutical agents. The present sequence represents a
 CC protein sequence of the invention.
 SQ Sequence 1004 AA;

Query Match 92.8%; Score 5077.5; DB 7; Length 1004;
 Best Local Similarity 91.9%; Pred. No. 0;
 Matches 976; Conservative 0; Mismatches 1; Indels 85; Gaps 3;
 QY 1 MLRTAGRDGLCRSLTYLEEAVALKKFKLVLTATLGECKIPWGSMEKAGPLEMAQLL 60
 DB 1 MLRTAGRDGLCRSLTYLEEAVALKKFKLVLTATLGECKIPWGSMEKAGPLEMAQLL 60
 QY 61 ITHFGPEAWRLALSTFERINRDLWERGQREDLVR----- 96
 DB 61 ITHFGPEAWRLALSTFERINRDLWERGQREDLVR----- 120
 QY 97 ---DPQETRYDYVRKFLMEDRNARLGEVNLSHRYTRLLLVKHSNPMQVQOQLDGTG 153

DB 121 PRKDPQETRYDYVRKFLMEDRNARLGEVNLSHRYTRLLLVKHSNPMQVQOQLDGTG 180
 QY 154 RGHARTVGHQASPIKIEITLPEPDERPEPRPTVYMQGAAGIGKSMIAHKVMDWADGKLF 213
 DB 181 RGHARTVGHQASPIKIEITLPEPDERPEPRPTVYMQGAAGIGKSMIAHKVMDWADGKLF 240
 QY 214 QGRFDYLYFYNCREMNSQATSCSMQDLIFSCWPSPSAPLOELIRVPERLLFIIDGFDL 273
 DB 241 QGRFDYLYFYNCREMNSQATSCSMQDLIFSCWPSPSAPLOELIRVPERLLFIIDGFDL 300
 QY 274 PSHPDQPGWCLCWEKRPTELINSLIRKKLPELSLITTRTALEKRLHLEHPRHV 333
 DB 301 PSHPDQPGWCLCWEKRPTELINSLIRKKLPELSLITTRTALEKRLHLEHPRHV 360
 QY 334 EILGFSEAEKKEYFYKYPHNAEQAGQVFNVDNEPLFTMCVFVLCVWVCTCQQOQLEG 393
 DB 361 EILGFSEAEKKEYFYKYPHNAEQAGQVFNVDNEPLFTMCVFVLCVWVCTCQQOQLEG 420
 QY 394 GGLLRQTSRTTAVYMLYLLSLMQPKPGAPRLQPPNQRGLCSLAADGLNQNKILFBEQD 453
 DB 421 GGLLRQTSRTTAVYMLYLLSLMQPKPGAPRLQPPNQRGLCSLAADGLNQNKILFBEQD 480
 QY 454 LRKHGLDGEDVSAPLANMNI FQKQINCERYYSFTHLSQEPFAAMYIILDEGEGAGDD 513
 DB 481 LRKHGLDGEDVSAPLANMNI FQKQINCERYYSFTHLSQEPFAAMYIILDEGEGAGDD 540
 QY 514 VTRLLTEYAFSERSFLALTSRFLGGLNEETRSLSKSLCKWSPHMKMDLLOWIQSKAQ 573
 DB 541 VTRLLTEYAFSERSFLALTSRFLGGLNEETRSLSKSLCKWSPHMKMDLLOWIQSKAQ 600
 QY 574 SDGSTLQOQSLEFPSCLYEIOEEFIQOALSHFQVIVVSNIAKMEHVSFCLKRCRSA 633
 DB 601 SDGSTLQOQSLEFPSCLYEIOEEFIQOALSHFQVIVVSNIAKMEHVSFCLKRCRSA 660
 QY 634 QVHLHYGATYSADGEDRARCAGAHLLVQL-PERTVLLDAYSHLAAALCTNPNLIELS 693
 DB 661 QVHLHYGATYSADGEDRARCAGAHLLVQL-PERTVLLDAYSHLAAALCTNPNLIELS 719
 QY 694 LYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSG 753
 DB 720 LYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSG 779
 QY 754 NGVGFPGMMLLCEGLRHPQCRLOMIQLRKQLESGACQEMASVLGTNPHLVLDLTGNAL 813
 DB 780 NGVGFPGMMLLCEGLRHPQCRLOMIQLRKQLESGACQEMASVLGTNPHLVLDLTGNAL 839
 QY 814 EDLGLRLICQGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSG 873
 DB 840 EDLGLRLICQGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSG 899
 QY 874 GVLLLCBGLRHPCKLQTLRLGICRLGSAACEGLSVVLQANHNLRELDLSFNDLGDWGLW 933
 DB 900 GVLLLCBGLRHPCKLQTLRL- 919
 QY 934 LLAEGLOHPACRLQKLVLDSCGLTAKACENLYFTLGINQTLTDLTLNNALGDTGVRLLC 993
 DB 961 LLAEGLOHPACRLQKLVLDSCGLTAKACENLYFTLGINQTLTDLTLNNALGDTGVRLLC 997

RESULT 10

ADP47728
 ID ADP47728 standard; protein; 950 AA.
 XX ADP47728;
 XX AC
 DT 12-AUG-2004 (first entry)
 XX Human Monarch-1 (CATERPILLER 19.3) isoform III protein.
 DE
 XX

KW Monarch-1; CATERPILLER 11.2; caspase recruitment domain;
KW CARD, transcription enhancer, R-binding, pyrin, lots of leucine repeat;
KW CATERPILLER 11.3; CATERPILLER 16.1; CATERPILLER 16.2; CIASI;
KW cold-induced autoinflammatory syndrome 1; antiinflammatory; cytostatic;
KW inflammatory disease; cancer; gene therapy; human; purine;
KW CATERPILLER 19.3; chromosome 19q13; multiple sclerosis; isoform III.
XX
OS Homo sapiens.
XX
XX WO2004034093-A2.
XX
XX 22-APR-2004.
XX
XX 30-APR-2003; 2003WO-US013562.
XX
XX 30-APR-2002; 2002US-0376626P.
XX
XX (UYNC-) UNIV NORTH CAROLINA.
XX
XX Ting JY, Linhoff MW, Harton JA, Williams KL, Lich J, O'Connor W;
XX Moore CB, Davis B, Brickey J, Conti B, Zhang J, Zhu X;
XX
XX WPI; 2004-348215/32.
XX
XX N-PSDB; ADP47727.
XX
XX New nucleic acid encoding Monarch-1, CATERPILLER 11.2, CATERPILLER 11.3,
XX CATERPILLER 16.1, CATERPILLER 16.2 or CIASI polypeptide, useful in
XX preparing a composition for treating inflammatory disease or cancer.
XX
XX Claim 18; SEQ ID NO 6; 205pp; English.
XX
XX The invention relates to a novel isolated nucleic acid encoding a Monarch
XX -1, CATERPILLER (CARD [caspase recruitment domain], transcription
XX enhancer, R(purine)-binding, pyrin, lots of leucine repeats) 11.2,
XX CATERPILLER 11.3, CATERPILLER 16.1, CATERPILLER 16.2 or CIASI (cold-
XX induced autoinflammatory syndrome 1) polypeptide comprising the amino
XX acids encoded by exon 6 and lacking the amino acids encoded by exon 4 or
XX its fragment. The nucleic acid of the invention demonstrates
XX antiinflammatory and cytostatic activities and may be useful in preparing
XX a composition for treating an inflammatory disease or cancer, possibly
XX via gene therapy. The current sequence is that of the human Monarch-1
XX (CATERPILLER 19.3) isoform III protein of the invention which is encoded
XX by DNA located on chromosome 19q13, in the multiple sclerosis
XX susceptibility region.
XX
XX Sequence 950 AA;
Query Match 87.7%; Score 4799.5; DB 8; Length 950;
Best Local Similarity 86.9%; Pred. No. 0;
Matches 923; Conservative 0; Mismatches 0; Indels 139; Gaps 2;

QY 1 MLRTAGRGCLRLSTYLEEAEVLEKFKFLYGTATLGECKIPWGSMEKAGPLEMAQLL 60
Db 1 MLRTAGRGCLRLSTYLEEAEVLEKFKFLYGTATLGECKIPWGSMEKAGPLEMAQLL 60

QY 61 ITHGPEAWRLALSTFERINRKNLWEGQREDLVR----- 96
Db 61 ITHGPEAWRLALSTFERINRKNLWEGQREDLVRTPPGPSSLGNQSTCLLEVLST 120

QY 97 ---DPQETRYDYVRKFRMEDNARLGCENLGHRYTRLLLVKHESNPMQVQQLDGT 153
Db 121 PRKDPQETRYDYVRKFRMEDNARLGCENLGHRYTRLLLVKHESNPMQVQQLDGT 180

QY 154 RGHARTVGHQASPIKIEITLFEPPERPEPPPTVMQAGIGKSLAHKWLMDWADGKLF 213
Db 181 RGHARTVGHQASPIKIEITLFEPPERPEPPPTVMQAGIGKSLAHKWLMDWADGKLF 240

QY 214 QRFDFYLFVINCENMQATSCMODLIFSCWPSPAPLOELIRVPERLLFIIDGFDELK 273
Db 241 QRFDFYLFVINCENMQATSCMODLIFSCWPSPAPLOELIRVPERLLFIIDGFDELK 300

QY 274 PSFHDPPQGWCLWEKRPTELLNSLIRKKLLPELSLLITRPTALEKHLRLEHPRHV 333
Db 274 PSFHDPPQGWCLWEKRPTELLNSLIRKKLLPELSLLITRPTALEKHLRLEHPRHV 333

Db 301 PSFHDPPQGWCLWEKRPTELLNSLIRKKLLPELSLLITRPTALEKHLRLEHPRHV 360
QY 334 EILGFSEARKEYFYKVFHNAEQAGVFNVDNEPLFTMCFVLVWVWVCTCLOOQLEG 393
Db 361 EILGFSEARKEYFYKVFHNAEQAGVFNVDNEPLFTMCFVLVWVWVCTCLOOQLEG 420
QY 394 GGLLRQTSRTTAVVYMLYLLSMQPKGAPRLQPPNQRGLCSLAADGLMNKILFEBQD 453
Db 421 GGLLRQTSRTTAVVYMLYLLSMQPKGAPRLQPPNQRGLCSLAADGLMNKILFEBQD 480
QY 454 LRKHGLDGEDVSAFLNMNIFOKDINCERYYSFIHLSFOEFFAAMYIILDEGGAGPDOD 513
Db 481 LRKHGLDGEDVSAFLNMNIFOKDINCERYYSFIHLSFOEFFAAMYIILDEGGAGPDOD 540
QY 514 VTRLITEYAFERSFLATLSRFLGLNEETRSHLEKSLCWKSPHISKMDLLQWISKAQ 573
Db 541 VTRLITEYAFERSFLATLSRFLGLNEETRSHLEKSLCWKSPHISKMDLLQWISKAQ 600
QY 574 SDGSTLQGSLEFFSCLYEIQEEFIQQAALSHFQVIVVSNIAASKMEHVMVSFCLKRCRSA 633
Db 601 SDGSTLQGSLEFFSCLYEIQEEFIQQAALSHFQVIVVSNIAASKMEHVMVSFCLKRCRSA 660
QY 634 QVHLHYGATYSADGEDRARCAGAHITLVQLRPERTVLLDAYSEHLAALCTNPNIELS 693
Db 661 QVHLHYGATYSADGEDRARCAGAHITLVQLRPERTVLLDAYSEHLAALCTNPNIELS 720
QY 694 LYRNALGSRGVKLLCOGLRHPNCKLQNLKRCRISSEACEDLSAALIANKNLTRMDLSG 753
Db 721 LYRNALGSRGVKLLCOGLRHPNCKLQNLKRCRISSEACEDLSAALIANKNLTRMDLSG 780
QY 754 NGVFPFMMLLCEGRHPQCRLOMTQLRKCOLESGACQEMASVLGTNPHLVELDTGNAL 813
Db 781 NGVFPFMMLLCEGRHPQCRLOMTQLRKCOLESGACQEMASVLGTNPHLVELDTGNAL 840
QY 814 EDLGLRLCOGLRHPVCRILRTLWLKICRLTAACDELASTLVNQSLRELDLSNELGDL 873
Db 841 EDLGLRLCOGLRHPVCRILRTLWL----- 863
QY 874 GVLLLCLEGRHPCTCKLQTLRLGICRLGSAACEGLSVLQANHLRELDLSFNDLGDWGLW 933
Db 864 ----- 863
QY 934 LLAEGLOHPACRLQKWLDSOGLTAKACENLYFTGINTLTDLYLTNNALGDTGVRLIC 993
Db 864 -----LWLDSCGLTAKACENLYFTGINTLTDLYLTNNALGDTGVRLIC 908
QY 994 KRLSHPGCKLRVLWLFGLMDLNKMTSHSLAALRVTKPYLDIGC 1035
Db 909 KRLSHPGCKLRVLWLFGLMDLNKMTSHSLAALRVTKPYLDIGC 950

RESULT 11
ADC31287
ID ADC31287 standard; protein; 865 AA.
XX
AC ADC31287;
XX
XX 18-DEC-2003 (first entry)
XX
XX Human novel polypeptide sequence, SEQ ID NO:1369.
XX
XX Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianemic; anticoagulant; thrombolytic; vulnery;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
XX gene therapy; chromosome 19.
XX
XX Homo sapiens.
XX
XX WO2003029271-A2.
PN

[illegible]

SQ Sequence 635 AA;
Query Match 61.5%; Score 3364; DB 8; Length 635;
Best Local Similarity 100.0%; Pred. No. 8.1e-313;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLRTAGDGLCRSLTYLEELAEVLEKFKLYLTATLGEGKIPWGSMEKAGPLEMAQLL 60
DB 1 MLRTAGDGLCRSLTYLEELAEVLEKFKLYLTATLGEGKIPWGSMEKAGPLEMAQLL 60
QY 61 ITHFGPEARLALSTPERINRKLWERGOREDLVRDPQETDYDYRRKFLMEDRNARL 120
DB 61 ITHFGPEARLALSTPERINRKLWERGOREDLVRDPQETDYDYRRKFLMEDRNARL 120
QY 121 GECVNLSHRYTRLLVKEHNPVQOQLDTRGHARTVGHQASPIKTIETLFEPPDEEP 180
DB 121 GECVNLSHRYTRLLVKEHNPVQOQLDTRGHARTVGHQASPIKTIETLFEPPDEEP 180
QY 181 EPRTVVMQGAAGIGKSMIAHKVMDWADGKLFQGRFDYLYINCREMNQSAATECSMDL 240
DB 181 EPRTVVMQGAAGIGKSMIAHKVMDWADGKLFQGRFDYLYINCREMNQSAATECSMDL 240
QY 241 IFSCWPPSAPLOBLIRVPERLLFIIDGFDLKPSPHDPQGWCLCWEKKRPTTELLNSL 300
DB 241 IFSCWPPSAPLOBLIRVPERLLFIIDGFDLKPSPHDPQGWCLCWEKKRPTTELLNSL 300
QY 301 IRKLLPELSLITRTPTALEKHLRLEHPRHVEILGFSEAEKKEFYKYFHNABQAGV 360
DB 301 IRKLLPELSLITRTPTALEKHLRLEHPRHVEILGFSEAEKKEFYKYFHNABQAGV 360
QY 361 FNYVRDNEPLFTMCVFPLVGVVCTCLQQLEGGLLRQTSRTTAVMYLYLLSLMQPKP 420
DB 361 FNYVRDNEPLFTMCVFPLVGVVCTCLQQLEGGLLRQTSRTTAVMYLYLLSLMQPKP 420
QY 421 GAPLQPPNQRGLCSLAADGLMNQKILFPEODLRKGLGDEVDVSAFLNNWIFQKDNCE 480
DB 421 GAPLQPPNQRGLCSLAADGLMNQKILFPEODLRKGLGDEVDVSAFLNNWIFQKDNCE 480
QY 481 RYYSFIHLSFOEFPAAMYIILDEGEGAGDPDQDVTLLTYVATERSFLALTISRFGLL 540
DB 481 RYYSFIHLSFOEFPAAMYIILDEGEGAGDPDQDVTLLTYVATERSFLALTISRFGLL 540
QY 541 NEETRSHLEKSLCWKSPHIMDLQWTSQAQSDGSTLQOGLSLEPSCLYIETQEEFIQ 600
DB 541 NEETRSHLEKSLCWKSPHIMDLQWTSQAQSDGSTLQOGLSLEPSCLYIETQEEFIQ 600
QY 601 QALSHFQVIVVSNIAKMEHNVSSFCILKCRSAQV 635
DB 601 QALSHFQVIVVSNIAKMEHNVSSFCILKCRSAQV 635

RESULT 15

ID ADE36457
AC ADE36457 standard; protein; 603 AA.

XX

AC ADE36457;

XX

XX 29-JAN-2004 (first entry)

DE Human PAN6 leucine-rich-repeat domain (LLR) seq id 64.

XX

XX

XX

XX

XX

XX

KW fibrosis; smooth muscle cell proliferation; balloon angioplasty;
KW restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
KW arthritis; lupus; schrojen's syndrome; Crohn's disease;
KW ulcerative colitis; graft versus host disease; stroke; heart failure;
KW neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
KW cancer therapy; PAAD domain family; human; PAN6; leucine-rich-repeat;
LLR.

XX

XX Homo sapiens.

XX US2003077699-A1.

PN 24-APR-2003.

XX

XX 25-SEP-2001; 2001US-00965621.

XX 26-SEP-2000; 2000US-00671760.

PR 26-SEP-2000; 2000US-0367367P.

XX

(REED/) REED J C.

(GODZ/) GODZIK A.

(CHUZ/) CHU Z.

(PAWL/) PAWLOWSKI K.

(FIOR/) FIORENTINO L.

(ARIZ/) ARIZA M E.

(STEH/) STEHLIK C.

XX

Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;

PI Stehlik C;

PI

XX WPI; 2002-471256/50.

XX

Novel isolated PAAD domain containing polypeptide useful for inducing

PT apoptosis by inhibiting nuclear factor kappa B activation and in gene

PT therapy for treating cancer.

XX

PS Claim 24; SEQ ID NO 64; 93pp; English.

XX

The invention describes an isolated PAAD domain containing polypeptide

CC (I) comprising 80% identity to the amino acid sequence of PAAD and

CC nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck

CC -like protein containing a caspase recruitment domain (ASC)-2 fully

CC defined in specification, where (I) is biologically active. (I) is useful

CC for identifying a (I)-associated polypeptide, an agent altering that

CC association and agents that modulate PAAD domain mediated inhibition of

CC nuclear factor kappa B (NFkappaB). A NB-ARC domain polypeptide is useful

CC for identifying an agent that modulates the activity of the NB-ARC domain

of (I). (I) or its functional fragments is useful in altering cellular or
CC biochemical process such as apoptosis, NFkappaB induction, cytokine
CC processing, cytokine receptor signaling caspase-mediated proteolysis or c
CC -Jun N-terminal kinase activation, thus having modulating effect on cell
CC life and death (apoptosis) inflammation, cell adhesion or other cellular
CC or biochemical processes. (I) is useful for treating cancer pathologies,
CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
CC hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
CC proliferation in arteries following balloon angioplasty (restenosis),
CC leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
CC lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft
CC versus host disease, stroke, heart failure, neurodegenerative diseases
CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
CC infection (HIV). (I) is useful for diagnosing cancer or monitoring cancer
CC therapy. This is the amino acid sequence of a human PAAD and nucleotide
CC binding protein PAN6 leucine-rich-repeat domain (LLR).

XX

Sequence 603 AA;

Query Match 57.8%; Score 3163; DB 5; Length 603;

Best Local Similarity 100.0%; Pred. No. 1.4e-293;

Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

QY 429 PNORGLCSLAADGLMNQKILFEBQDLRKGLGDEVDVSAFLNNWIFQKDNCEYYSFIHL 488

DB 1 PNORGLCSLAADGLMNQKILFEBQDLRKGLGDEVDVSAFLNNWIFQKDNCEYYSFIHL 60

Qy 489 SFQEFFAAMYIILDEGEGAGDPQDVTRLLTEYAFSERSFLALTSRPLFGLNEETRSHL 548
| | | | |
Db 61 SFQEFFAAMYIILDEGEGAGDPQDVTRLLTEYAFSERSFLALTSRPLFGLNEETRSHL 120
| | | | |
Qy 549 EKSLCKVSPHIMDQLQWIOSKQSDGSTLQOQSLEFFSCLYEIQEEFFIQOALSHPQV 608
| | | | |
Db 121 EKSLCKVSPHIMDQLQWIOSKQSDGSTLQOQSLEFFSCLYEIQEEFFIQOALSHPQV 180
| | | | |
Qy 609 IVVSNIAKMEHMYSSFCCLKCRSAQVHLHYGATYSADGEDRARCAGAHILLVQLRPER 668
| | | | |
Db 181 IVVSNIAKMEHMYSSFCCLKCRSAQVHLHYGATYSADGEDRARCAGAHILLVQLRPER 240
| | | | |
Qy 669 TVLLDAYSEHLAAALCTNPNIIEISLYRNALGSRGVKLLCOGLRHPNCKLQNLKRCRI 728
| | | | |
Db 241 TVLLDAYSEHLAAALCTNPNIIEISLYRNALGSRGVKLLCOGLRHPNCKLQNLKRCRI 300
| | | | |
Qy 729 SSSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLICEGLRHPQCRLOMIQLRKCQLESG 788
| | | | |
Db 301 SSSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLICEGLRHPQCRLOMIQLRKCQLESG 360
| | | | |
Qy 789 ACQEMASVLGTNPHLVELDLTGNALEDLGLRLLCOGLRHPVCRLLTLWKICRLTAAACD 848
| | | | |
Db 361 ACQEMASVLGTNPHLVELDLTGNALEDLGLRLLCOGLRHPVCRLLTLWKICRLTAAACD 420
| | | | |
Qy 849 ELASTLSVNQSLRELDLSLNEIGDLGVLLLCGLRHPPTCKLQTLRLGICRLGSAACEGLS 908
| | | | |
Db 421 ELASTLSVNQSLRELDLSLNEIGDLGVLLLCGLRHPPTCKLQTLRLGICRLGSAACEGLS 480
| | | | |
Qy 909 VVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKLDSCGLTAKACENLYFTL 968
| | | | |
Db 481 VVLQANHNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKLDSCGLTAKACENLYFTL 540
| | | | |
Qy 969 GINQTLTDLYLTNNALGDTGVRLLCRLSHPGCKLRVLWLFQMDLNKMTSHSLAALRVTK 1028
| | | | |
Db 541 GINQTLTDLYLTNNALGDTGVRLLCRLSHPGCKLRVLWLFQMDLNKMTSHSLAALRVTK 600
| | | | |
Qy 1029 PYL 1031
| | |
Db 601 PYL 603

Search completed: February 8, 2005, 14:49:57
Job time : 180 secs

This Page Blank (uspto)